ITUPublications

International Telecommunication Union
Telecommunication Standardization Sector

# AI for Good Innovate for Impact Report 2025

## Disclaimer

The views expressed in this publication are those of the authors and do not necessarily reflect the views of ITU. Any references made to specific countries, companies, products, initiatives or guidelines do not in any way imply that they are endorsed or recommended by ITU, the authors, in preference to others of similar nature that are not mentioned. Requests to reproduce extracts of this publication may be submitted to *jur@itu.int*. This document is intended for informational purposes only. Information provided is correct as of June 2025.

The designations employed and the presentation of the material in this publication do not imply the expression of any opinion whatsoever on the part of ITU concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries.

The mention of specific companies or certain manufacturer products does not imply that they are endorsed or recommended by ITU in preference to others of a similar nature that are not mentioned.

All reasonable precautions have been taken by ITU to verify the information contained in this publication. However, the published material is being distributed without warranty of any kind, either expressed or implied. The responsibility for the interpretation and use of the material lies with the reader.

The opinions, findings and conclusions expressed in this publication do not necessarily reflect the views of ITU or its membership.


## ISBN

978-92-61-40961-6

# AI for Good Innovate for Impact

Report 2025

# Foreword (TSB Director)

Artificial intelligence (AI) is reshaping the way we address complex societal challenges, offering new possibilities in areas such as healthcare, climate resilience, education, and digital inclusion. The Innovate for Impact project was launched in 2024 to identify, support, and showcase practical AI solutions that bring tangible benefits to people and communities. The project brings together contributions from multiple sectors—industry, academia, government, civil society, and UN organizations—creating a collaborative environment where ideas, knowledge, and experience are shared to accelerate responsible and impactful AI deployment.

A key part of the initiative involves the sharing of use cases, impactful AI applications and global lessons and expertise from around the world. In 2025, building on the success of its first edition, the project expanded its scope through an open call for AI use cases and AI Scholars. We received 234 use case submissions from 32 countries, out of which 160 were selected for inclusion in this final report. These use cases span eleven key domains and reflect both the diversity of global innovation, regional solutions with lessons learnt and the practical ways in which AI is being applied to solve real-world problems.

To enhance the quality and relevance of the evaluation process, we established a Technical Advisory Committee (TAC). From 89 expert nominations received from 32 countries, 32 members were carefully selected based on their expertise, regional representation, and commitment to responsible AI. The TAC brings together specialists from AI research, engineering, ethics, social sciences, and development policy, and has played a central role in reviewing submissions and shaping the strategic direction of the project.

We also launched a call for AI Scholars to support the initiative through research and engagement. From 171 applications submitted by candidates from 37 countries, 16 outstanding scholars were selected. They worked closely with use case authors to analyse, refine, and document the AI solutions, contributing and imbibing valuable insights and building bridges between communities and domains such as research, policy, and practice.

The 2024 study focused on collecting and analysing concept-based use cases. These initial examples offered insights into potential AI applications across diverse domains. In 2025, we go beyond concept based use cases, to evaluate practical pilots and real-world implementations. Validating functionality with real world experiences and gather user feedback to refine solutions and demonstrate feasibility at scale, was the focus in 2025. Going forward in 2026 - Interoperability and scaling would be the logical next focus area. By encouraging knowledge sharing and common frameworks, this phase will help translate individual successes into adaptable models that can be widely applied across regions and sectors.

We are grateful to all who contributed to this report, especially the use case authors, scholars, and TAC members for their commitment and insight. We also thank China's Ministry of Industry and Information Technology for its generous support of this project.

The Innovate for Impact project has matured over the last 2 years of our journey and now, we look forward to deepening our collaboration with partners worldwide. This will involve moving towards latest technologies such as AI Agents, and study of implementation experiences in the coming years.

Seizo Onoe
Director of the Telecommunication Standardization Bureau
International Telecommunication Union

# Table of contents


Foreword (TSB Director) ..... ii

1. Introduction ..... 1

2. Project introduction and selection process ..... 2

2.1 AI use cases ..... 2

2.2 Technical Advisory Committee formation and Selection ..... 5

2.3 Scholars Selection ..... 7

3. Learnings and conclusion ..... 8

3.1. Cross-Sector Insights from Global Use Cases ..... 8

3.2. Key Insights from 2025 Innovate for Impact analysis ..... 9

3.3. Future work ..... 10

4. Innovate for impact use cases ..... 11

4.1. Category 1: Healthcare ..... 11

Use Case 1: Intelligent Applications Based on Medical Large Models to Enhance People's Healthcare Experience and Health Status ..... 11

1 Use Case Summary Table ..... 11

2 Use Case Description ..... 12

3 Use Case Requirements ..... 15

4 Sequence Diagram ..... 16

5 References ..... 16

Use Case 2: AI based Virtual Yoga Instructor for Good health ..... 18

1 Use Case Summary Table ..... 18

2 Use Case Description ..... 19

3 Use Case Requirements ..... 21

4 Sequence Diagram ..... 21

5 References ..... 22

Use Case 3: CAFEIN – Federated Learning Platform ..... 23

1 Use Case Summary Table ..... 23

2 Use Case Description ..... 24

3 Use Case Requirements ..... 25

4 Sequence Diagram ..... 26
5 References ..... 27

Use Case 4: Smartphone-over-Microscope Diagnosis: Empowering Community Health Workers with TinyML for Malaria Parasite Detection in Rural Areas ..... 28

1 Use Case Summary Table ..... 28
2 Use Case Description ..... 29
3 Use Case Requirements ..... 31
4 Sequence Diagram ..... 32
5 References ..... 32

Use Case 5: AI-Driven Biometric Animal Registration for Animal Welfare and Global Health ..... 33

1 Use case summary table ..... 33
2 Use case description ..... 34
3 Use case requirements ..... 35
4 Sequence diagram ..... 36
5 References ..... 36

Use Case 6: AI-Powered Dental CAD Solution for Fully Automated Digital Prosthesis Design ..... 37

1 Use Case Summary Table ..... 37
2 Use Case Description ..... 38
3 Use Case Requirements ..... 40
4 Sequence Diagram ..... 40
5 References ..... 40

Use Case 7: AI-Powered Smart Healthcare for Remote & Underserved Communities ..... 42

1 Use Case Summary Table ..... 42
2 Use Case Description ..... 43
3 Use Case Requirements ..... 44
4 Sequence Diagram ..... 45
5 References ..... 46

Use Case 8: Geospatial Mapping of Otitis Media (OM) and Social Determinants in Remote Northern Territory Communities ..... 47

1 Use case summary Table ..... 47
2 Use Case Description: ..... 48
3 Use case Requirements ..... 50

4 Sequence Diagram: ..... 50
5 References ..... 50

Use Case 9: AI Senior Care Contents: Ensuring Joy and Cognitive Enhancement for Seniors ..... 52

1 Use Case Summary Table ..... 52
2 Use Case Description ..... 54
3 Use Case Requirements ..... 55
4 Sequence Diagram ..... 56
5 References ..... 57

Use Case 10: AI-mediated Interactive Health Messaging for Community Health Promotion in Low-Middle Income Countries. ..... 58

1 Use Case Summary Table ..... 58
2 Use Case Description ..... 59
3 Use Case Requirements ..... 61
4 Sequence Diagram ..... 62
5 References ..... 62

Use Case 11: Necrosis Detection in Meningioma Images ..... 64

1 Use Case Summary Table ..... 64
2 Use Case Description ..... 65
3 Use Case Requirements ..... 67
4 Sequence Diagram ..... 68
5 References ..... 68

Use Case 12: AI-Based Horse Identification and Gait Analysis ..... 69

1 Use Case Summary Table ..... 69
2 Use Case Description ..... 70
3 Use Case Requirements ..... 71
4 Sequence Diagram ..... 72
5 References ..... 72

Use Case 13: AI-Driven Optimization of Clinical Trial Preparation: From Protocol Design to CRO Selection ..... 73

1 Use Case Summary Table ..... 73
2 Use Case Description ..... 74
3 Use Case Requirements ..... 75
4 Sequence Diagram ..... 75

5 References ... 76

Use Case 14: AI-Enabled Pathology Model for Precision Diagnosis and Treatment ... 78

1 Use Case Summary Table ... 78
2 Use Case Description ... 79
3 Use Case Requirements ... 83
4 Sequence Diagram ... 83
5 References ... 84

Use Case 15: AI Specialist Diagnosis and Treatment Resources - Ophthalmology Model ... 85

1 Use Case Summary Table ... 85
2 Use Case Description ... 86
3 Use Case Requirements ... 89
4 Sequence Diagram ... 90
5 References ... 90

Use Case 16: Traditional Chinese Medicine Industry Leigong Large Model ... 91

1 Use Case Summary Table ... 91
2 Use Case Description ... 93
3 Use Case Requirements ... 94
4 Sequence Diagram ... 95
5 References ... 96

Use Case 17: MoralQuest: AI-Powered Gamification of Indian Moral Tales for Neurodiverse Social Development in children with Autism Spectrum Disorder (ASD) ... 97

1 Use Case Summary Table ... 97
2 Use Case Description ... 99
3 Use Case Requirements ... 100
4 Sequence Diagram ... 102
5 References ... 102

Use Case 18: Smart Rehabilitation: AI-Driven Exercise Assessment and Live Feedback for Enhanced Recovery ... 104

1 Use Case Summary Table ... 104
2 Use Case Description ... 105
3 Use Case Requirements ... 107
4 Sequence Diagram ... 107
5 References ... 108

Use Case 19: AI Powered Webisite of UAE Government Healthcare System ..... 109

1 Use Case Summary Table ..... 109
2 Use Case Description..... 112
3 Use case Requirements ..... 113
4 Sequence Diagram ..... 114
5 References..... 115

Use Case 20: An AI-Powered Augmentative and Alternative Communication (AAC) tool ..... 116

1 Use Case Summary Table ..... 116
2 Use Case Description..... 117
3 Use Case Requirements ..... 118
4 Sequence Diagram ..... 119
5 Reference ..... 119

Use Case 21: AI-Based Health Monitoring and Advisory System..... 120

1 Use Case Summary Table ..... 120
2 Use Case Description..... 120
3 Use Case Requirements ..... 121
4 Sequence Diagram ..... 121
5 References..... 122

Use Case 22: EquiDermAI: An Inclusive Deep-Generative Framework to Overcome Racial & Regional Barriers in Skin Disease Diagnosis..... 123

1 Use case summary table..... 123
2 Use case description..... 125
3 Use case requirements..... 127
4 Sequence diagram..... 127
5 References..... 128

Use Case 23: Empowering Student Mental Well-being with AI: Intelligent Journaling, Emotional Insights, and Personalized Growth..... 129

1 Use case summary table..... 129
2 Use Case Description: ..... 130
3 Use case requirements..... 131
4 Sequence diagram..... 131
5 References..... 132

Use Case 24: Brain-wave driven wheelchair for paralysis patients using EEG and ML..... 133

1 Use Case Summary Table ..... 133
2 Use Case Description ..... 134
3 Use Case Requirements ..... 135
4 Sequence Diagram ..... 136
5 References ..... 136

Use Case 25: Smart Watch based Integration System for Habituating Face Paralysis Monitoring ..... 138

1 Use Case Summary Table ..... 138
2 Use Case Description ..... 139
3 Use Case Requirements ..... 140
4 Sequence Diagram ..... 141
5 References ..... 141

Use Case 26: AI-Enhanced Organoid Platform for Pandemic Preparedness and Personalized Medicine ..... 142

1 Use Case Summary Table ..... 142
2 Use Case Summary Table ..... 143
3 Use case requirements ..... 144
4 Sequence Diagram Design ..... 145
5 References ..... 146

Use Case 27: Multi-Cancer Early Screening with Non-contrast CT and Deep Learning. ... 147

1 Use Case Summary Table ..... 147
2 Use Case Description ..... 148
3 Use Case Requirements ..... 149
4 Sequence Diagram ..... 150
5 References ..... 150

4.2. Category 2: Climate Change ..... 152

Use Case 1: GeoGPT: AI-driven Marine Ecosystem Monitoring and Protection ..... 152

1 Use Case Summary Table ..... 152
2 Use Case Description ..... 153
3 Use Case Requirements ..... 155
4 Sequence Diagram ..... 156
5 References ..... 156

Use Case 2: AI-Powered High-Precision Weather Forecasting Model ..... 157

1 Use Case Summary Table ..... 157
2 Use Case Description ..... 158
3 Use Case Requirements ..... 159
4 Sequence Diagram ..... 160
5 References ..... 160

Use Case 3: AI-powered Monitoring of Mangrove Biomass and Carbon Stocks ..... 161

1 Use Case Summary Table ..... 161
2 Use Case Description ..... 162
3 Use Case Requirements ..... 163
4 Sequence Diagram ..... 163
5 References ..... 163

Use Case 4: AIrth – AI for Energy Efficiency and Emissions Reduction ..... 164

1 Use Case Summary Table ..... 164
2 Use Case Description ..... 165
3 Use Case Requirements ..... 166
4 Sequence Diagram ..... 166
5 References ..... 166

Use Case 5: LLM-driven creation of natural hazard geodatabase from digital mass media ..... 167

1 Use Case Summary Table ..... 167
2 Use Case Description ..... 168
3 Use Case Requirements ..... 168
4 Sequence Diagram ..... 169
5 References ..... 169

Use Case 6: AI-powered Precision Recycling System (WIMPLE) ..... 170

1 Use Case Summary Table ..... 170
2 Use Case Description ..... 171
3 Use Case Requirements ..... 176
4 Sequence Diagram ..... 177
5 References ..... 177

Use Case 7: LLM empowered smart hydrology ..... 178

1 Use Case Summary Table ..... 178
2 Use Case Description ..... 178

3 Use Case Requirements ..... 180
4 Sequence Diagram ..... 181
5 References ..... 181

Use Case 8: AI for green multi-cluster: Intelligent management towards green and low-carbon, large-scale multi-clusters ..... 182

1 Use Case Summary Table ..... 182
2 Use Case Description ..... 183
3 Use Case Requirements ..... 185
4 Sequence Diagram ..... 186
5 References ..... 186

Use Case 9: Combating wildlife trade with an advanced AI solution: a collaboration between Baidu PaddlePaddle and IFAW ..... 188

1 Use Case Summary Table ..... 188
2 Use Case Description ..... 188
3 Use Case Requirements ..... 189
4 Sequence Diagram ( Tobe updated) ..... 190
5 References ..... 190

Use Case 10: Quartz Solar AI Cloudcasting - predicting cloud movements through satellite imagery ..... 191

1 Use Case Summary Table ..... 191
2 Use Case Description ..... 192
3 Use Case Requirements ..... 192
4 Sequence Diagram ..... 193
5 References ..... 193

Use Case 11: AI biodiversity monitoring ..... 194

1 Use Case Summary Table ..... 194
2 Use Case Description ..... 194
3 Use Case Requirements ..... 195
4 Sequence Diagram ..... 195
5 References ..... 196

Use Case 12: AI for Weather Prediction for Africa (Project Umoya) ..... 197

1 Use Case Summary Table ..... 197
2 Use Case Description ..... 197
3 Use Case Requirements ..... 198

4 Sequence Diagram ..... 199
5 References ..... 199

Use Case 13: Analysing Impact of Regional Weather on PLC-based Data Transmission in Malaysia, and Nigeria. ..... 200

1 Use Case Summary Table ..... 200
2 Use Case Description ..... 201
3 Use Case Requirements ..... 202
4 Sequence Diagram ..... 202
5 References ..... 203

Use Case 14: Remote Visual Tracking Using Semantic Scene Understanding and GenAI-based Predictive Rendering Under Disaster Situation ..... 204

1 Use Case Summary Table ..... 204
2 Use Case Description ..... 205
3 Use Case Requirements ..... 206
4 Sequence Diagram ..... 207
5 References ..... 207

Use Case 15: AI Enables New Energy Meteorological Power Forecast ..... 208

1 Use Case Summary Table ..... 208
2 Use Case Description ..... 209
3 Use Case Requirements ..... 211
4 Sequence Diagram ..... 212
5 References ..... 212

Use Case 16: AI-Driven Early Warning System for Oceanic Dead Zones ..... 213

1 Use Case Summary Table ..... 213
2 Use Case Description ..... 214
3 Use Case Requirements ..... 216
4 Sequence Diagram ..... 217
5 References ..... 218

Use Case 17: AI-Powered RF Signal Locator for Advanced Search and Rescue Operations ..... 219

1 Use Case Summary Table ..... 219
2 Use Case Description ..... 219
3 Use Case Requirements ..... 220
4 Sequence Diagram ..... 220

5 References ..... 221

Use Case 18: Cooling the Concrete Jungle: An AI-Enabled Urban Heat Island Mitigation Advisor for Sustainable Development. ..... 222

1 Use Case Summary Table ..... 222
2 Use Case Description ..... 223
3 Use Case Requirements ..... 224
4 Sequence Diagram and dashboard ..... 224
5 References ..... 225

Use Case 19: Spatio-Temporal Microclimate Prediction using Deep Learning ..... 226

1 Use Case Summary Table ..... 226
2 Use Case Description ..... 227
3 Use Case Requirements ..... 228
4 Sequence Diagram ..... 228
5 References ..... 229

Use Case 20: AI-Based Real-Time Elephant Monitoring and Human Elephant Conflict Mitigation System ..... 231

1 Use Case Summary Table ..... 231
2 Use Case Description ..... 232
3 Use Case Requirements ..... 233
4 Sequence Diagram ..... 233
5 References ..... 234

Use Case 21: Driven Meteorological Modeling: Transforming Wind Speed and Direction Prediction ..... 235

1 Use Case Summary Table ..... 235
2 Use Case Description ..... 236
3 Use Case Requirements ..... 237
4 Sequence Diagram ..... 237
5 References ..... 238

Use Case 22: Polar Eco-Carbon Intelligence Platform (PECIP): an AI-Driven ESG Project for Climate Action and Global Sustainability ..... 239

1 Use Case Summary Table ..... 239
2 Use Case Description ..... 241
3 Use Case Requirements ..... 246
4 Sequence Diagram ..... 247

5 References ..... 249

Use Case 23: AI-Powered Chatbot for Disaster Response ..... 250

1 Use Case Summary Table ..... 250
2 Use Case Description ..... 250
3 Use Case Requirements ..... 251
4 Sequence Diagram ..... 252
5 References ..... 252

4.3. Category 3: 5G ..... 253

Use Case 1: AI Full-Process Low-Code Middle Platform Based on the '1+1+3+N' Architecture ..... 253

1 Use Case Summary Table ..... 253
2 Use Case Description ..... 254
3 Use Case Requirements ..... 257
4 Sequence Diagram ..... 258
5 References ..... 258

Use Case 2: Empowering cloud network operation with self-intelligence through the network large model toolchain platform ..... 260

1 Use Case Summary Table ..... 260
2 Use Case Description ..... 261
3 Use Case Requirements ..... 262
4 Sequence Diagram ..... 263
5 References ..... 263

Use Case 3: AI-Driven Digital Twins for Energy-Efficient Telco Networks ..... 265

1 Use Case Summary Table ..... 265
2 Use Case Description ..... 266
3 Use Case Requirements ..... 270
4 Sequence Diagram ..... 270
5 References ..... 271

Use Case 4: Integrated AI and Wireless for Sustainable Mobile Network Evolutions ..... 272

1 Use Case Summary Table ..... 272
2 Use Case Description ..... 273
3 Use Case Requirements ..... 275
4 Sequence Diagram ..... 276

5 References ..... 276

Use Case 5: Generative Artificial Intelligence Content Security Detection and Protection Platform ..... 278

1 Use Case Summary Table ..... 278
2 Use Case Description ..... 279
3 Use Case Requirements ..... 280
4 Sequence Diagram ..... 280
5 References ..... 281

Use Case 6: The "CloudTop": AI-driven Marketing and Service Intelligentized Applications of China Telecom Cloud ..... 282

1 Use Case Summary Table ..... 282
2 Use Case Description ..... 283
3 Use Case Requirements ..... 286
4 Sequence Diagram ..... 287
5 References ..... 288

Use Case 7: “Di Ting” Customer Service Agent ..... 289

1 Use Case Summary Table ..... 289
2 Use Case Description ..... 290
3 Use Case Requirements ..... 292
4 Sequence Diagram ..... 293
5 References ..... 293

Use Case 8: Intelligent Agent Framework and Practice for Network Operations and Maintenance Based on the Collaboration of Large and Small Models ..... 294

1 Use Case Summary Table ..... 294
2 Use Case Description ..... 295
3 Use Case Requirements ..... 297
4 Sequence Diagram ..... 298
5 References ..... 298

Use Case 9: Smart Assistant for Telecom Business Acceptance Based on the “Hui Xin” Large Language Model ..... 299

1 Use Case Summary Table ..... 299
2 Use Case Description ..... 300
3 Use Case Requirements ..... 302
4 Sequence Diagram ..... 302

5 References ... 302

Use Case 10: AI-Driven RF-Based Object Detection and Classification for 5G-Advanced and Beyond ... 303

1 Use Case Summary Table ... 303
2 Use Case Description ... 304
3 Use Case Requirements ... 307
4 Sequence Diagram ... 308
5 References ... 308

Use Case 11: AI-Assisted Smart and Efficient Radio Access Networks ... 309

1 Use Case Summary Table ... 309
2 Use Case Description ... 310
3 Use Case Requirements ... 312
4 Sequence Diagram ... 313
5 References ... 313

Use Case 12: Dynamic Instance Switching in 5G Network Slices Based on Load ... 314

1 Use Case Summary Table ... 314
2 Use Case Description ... 315
3 Use Case Requirements ... 316
4 Sequence Diagram ... 316
5 References ... 317

Use Case 13: Intelligent Symbiosis: AI + 5G-A Redefining Sustainable Guardianship in Human-Wildlife Shared Zones ... 318

1 Use Case Summary Table ... 318
2 Use Case Description ... 318
3 Use Case Requirements ... 320
4 Sequence Diagram ... 321
5 References ... 321

Use Case 14: AI-Powered Network O&M, Ensuring Uninterrupted Network During Power Rationing ... 322

1 Use Case Summary Table ... 322
2 Use Case Description ... 323
3 Use Case Requirements ... 324
4 Sequence Diagram ... 325
5 References ... 325

Use Case 15: Digital green transformation of network infrastructure based on artificial intelligence ..... 327

1 Use Case Summary Table ..... 327
2 Use Case Description ..... 328
3 Use Case Requirements ..... 330
4 Sequence Diagram ..... 331
5 References ..... 331

Use Case 16: AI-Driven Predictive Beamforming for Climate-Resilient 5G/6G Networks ..... 333

1 Use Case Summary Table ..... 333
2 Use Case Description ..... 334
3 Use Case Requirements ..... 336
4 Sequence Diagram ..... 336
5 References ..... 337

Use Case 17: Anomaly Detection Across 5G Slices Using Federated Learning ..... 338

1 Use Case Summary Table ..... 338
2 Use Case Description ..... 339
3 Use Case Requirements ..... 340
4 Sequence Diagram ..... 340
5 References ..... 341

Use Case 18: GIS-Integrated Digital Twin for 5G/6G Energy Efficient Base Station ..... 342

1 Use Case Summary Table ..... 342
2 Use Case Description ..... 343
3 Use Case Requirements ..... 345
4 Sequence Diagram ..... 346
5 References ..... 346

Use Case 19: Bandwidth efficient live interaction with virtual 3D demonstrator using semantic communication and GenAI ..... 347

1 Use Case Summary Table ..... 347
2 Use Case Description ..... 348
3 Use Case Requirements ..... 350
4 Sequence Diagram ..... 351
5 References ..... 351

Use Case 20: AI-Driven Free Space Optical Communication System ..... 353

1 Use Case Summary Table ..... 353
2 Use Case Description ..... 354
3 Use Case Requirements ..... 356
4 Sequence Diagram ..... 356
5 References ..... 356

4.4. Category 4: Productivity ..... 357

Use Case 1: Non-Profit Organizations – AI Powered Registration Process ..... 357

1 Use Case Summary Table ..... 357
2 Use Case Description ..... 357
3 Use Case Requirements ..... 360
4 Sequence Diagram ..... 360
5 References ..... 363

Use Case 2: AI-Powered GaP Repository: Democratizing Access to Political Knowledge for Young Women in South Asia ..... 364

1 Use Case Summary Table ..... 364
2 Use Case Description ..... 365
3 Use Case Requirements ..... 366
4 Sequence Diagram ..... 367
5 References ..... 367

Use Case 3: HEART Engine is an AI-Powered Public Interest Tech Certification Platform to Build Future AI Public Interest Workforce ..... 368

1 Use Case Summary Table ..... 368
2 Use Case Description ..... 372
3 Use Case Requirements ..... 373
4 Sequence Diagram ..... 373
5 References ..... 374

Use Case 4: Making Everyone the Director of Their Own Life ..... 375

1 Use Case Summary Table ..... 375
2 Use Case Description ..... 375
3 Use Case Requirements ..... 377
4 Sequence Diagram ..... 378
5 References ..... 378

Use Case 5: Radiolytica - Leveraging AI Tools to Protect Communities from Harmful Content in Eastern DRC ..... 379

1 Use Case Summary Table ..... 379
2 Use Case Description..... 380
3 Use Case Requirements ..... 382
4 Sequence Diagram ..... 383
5 References..... 383

Use Case 6: Artificial Intelligence-Generated Content Moderation ..... 384

1 Use Case Summary Table ..... 384
2 Use Case Description..... 384
4 Sequence Diagram ..... 388
5 References..... 389

Use Case 7: AI model to summarize, rate, and flag potentially harmful clauses found in terms of service or privacy policy agreements for digital services ..... 390

1 Use Case Summary Table ..... 390
2 Use Case Description..... 390
3 Use Case Requirements ..... 392
4 Sequence Diagram ..... 392
5 References..... 393

Use Case 8: Uncovering Delays: Leveraging Data Science and AI to Streamline Refugee Status Determination Procedures ..... 394

1 Use Case Summary Table ..... 394
2 Use Case Description..... 396
3 Use Case Requirements ..... 399
4 Sequence Diagram ..... 400
5 References..... 400

Use Case 9: Developing an AI-powered translation app to improve workplace safety at construction and shipbuilding worksites..... 402

1 Use Case Summary Table ..... 402
2 Use Case Description..... 403
3 Use Case Requirements ..... 406
4 Sequence Diagram ..... 407
5 References..... 407

Use Case 10: FAHR Smart App - AI-powered HR Solutions for Government Employees . 408

1 Use Case Summary Table ..... 408
2 Use Case Description..... 409

3 Use Case Requirements ..... 410
4 Sequence Diagram ..... 410
5 References ..... 410

Use Case 11: Virtual Legal Assistant of the Rights Mapping and Analysis Platform (RiMAP) to Improve Legal Analysis and Access to Rights and Solutions for Forcibly Displaced and Stateless Persons ..... 411

1 Use Case Summary Table ..... 411
2 Use Case Description ..... 412
3 Use Case Requirements ..... 415
4 Sequence Diagram ..... 416
5 References ..... 416

Use Case 12 : Implementation Framework for AI-Enabled HRM in UAE Limited Liability Companies ..... 417

1 Use Case Summary Table ..... 417
2 Use Case Description ..... 418
3 Use Case Requirements ..... 419
4 Sequence Diagram ..... 419
5 References ..... 419

Use Case 13: Lost in Translation? AI's Promise and Peril for Media in Low-Resource Languages ..... 420

1 Use Case Summary Table ..... 420
2 Use Case Description ..... 421
3 Use Case Requirements ..... 421
4 Sequence Diagram ..... 422
5 References ..... 422

Use Case 14: AI-Driven Wireless Audio Conference System for Smartphone ..... 423

1 Use Case Summary Table ..... 423
2 Use Case Description ..... 424
3 Use Case Requirements ..... 425
4 Sequence Diagram ..... 426
5 Reference ..... 426

Use Case 15: AI-Powered Universal Communication Platform: Breaking Down Barriers Through Collaborative Innovation ..... 427

1 Use Case Summary Table ..... 427

2 Use Case Description ..... 428
3 Use Case Requirements ..... 431
4 Sequence Diagram ..... 432
5 References ..... 433

Use Case 16: AI Cyber Security Edge and Cloud based Intrusion Detection System for IoT ..... 435

1 Use Case Summary Table ..... 435
2 Use Case Description ..... 436
3 Use Case Requirements ..... 437
4 Sequence Diagram ..... 438
5 References ..... 438

Use Case 17: The future of Work and Attached Intelligence for Knowledge Workers: A Framework for Cognitive Resilience and Mindful Productivity ..... 439

1 Use Case Summary Table ..... 439
2 Use Case Description ..... 440
3 Use Case Requirements ..... 442
4 Sequence Diagram ..... 444
5 References ..... 444

Use Case 18: AI-Powered Talent Matching Platform ..... 445

1 Use Case Summary Table ..... 445
2 Use Case Description ..... 445
3 Use Case Requirements ..... 446
4 Sequence Diagram ..... 446
5 References ..... 446

4.5. Category 5: Manufacturing ..... 447

Use Case 1: Morphing-BOT: Revolutionizing Global Infrastructure with AI-Powered Smart Inspection and Predictive Maintenance ..... 447

1 Use Case Summary Table ..... 447
2 Use Case Description ..... 448
3 Use Case Requirements ..... 454
4 Sequence Diagram ..... 455
5 References ..... 456

Use Case 2: AI-based Semiconductor Design Automation and Optimization ..... 457

1 Use Case Summary Table ..... 457

2 Use Case Description ..... 458
3 Use Case Requirements ..... 460
4 Sequence Diagram ..... 460
5 References ..... 461

Use Case 3: Intelligent Transmission Section Flow Control Agent in Power Systems ..... 462

1 Use Case Summary Table ..... 462
2 Use Case Description ..... 463
3 Use Case Requirements ..... 465
4 Sequence Diagram ..... 466
5 References ..... 466

Use Case 4: AI-Driven Virtual Try-On and Garment Generation System for Sustainable Indian Fashion ..... 467

1 Use Case Summary Table ..... 467
2 Use Case Description ..... 468
3 Use Case Requirements ..... 471
4 Sequence Diagram ..... 471
5 References ..... 471

Use Case 5: Explosion-Proof Intelligent Inspection Robot Solution ..... 473

1 Use Case Summary Table ..... 473
2 Use Case Description ..... 474
3 Use Case Requirements ..... 476
4 Sequence Diagram ..... 476
5 References ..... 476

Use Case 6: Generative AI-Driven Full-Process Computer-Aided Design ..... 477

1 Use Case Summary Table ..... 477
2 Use Case Description ..... 477
3 Use Case Requirements ..... 479
4 Sequence Diagram ..... 479
5 References ..... 480

Use Case 7: AI Enables Intelligent Manufacturing of the Automobile Industry ..... 481

1 Use Case Summary Table ..... 481
2 Use Case Description ..... 482
3 Use Case Requirements ..... 484

4 Sequence Diagram ..... 485
5 References ..... 485
Use Case 8: The Artificial Intelligence Robot for Cement Testing ..... 486
1 Use Case Summary Table ..... 486
2 Use Case Description ..... 487
3 Use Case Requirements ..... 488
4 Sequence Diagram ..... 489
5 References ..... 489
Use Case 9: Synergy of Large Model and Small Model for Port Safety Management ..... 490
1 Use Case Summary Table ..... 490
2 Use Case Description ..... 491
3 Use Case Requirements ..... 493
4 Sequence Diagram ..... 494
5 References ..... 494
Use Case 10: AI-Based eXtended Reality Training for Automation ..... 495
1 Use Case Summary Table ..... 495
2 Use Case Description ..... 496
3 Use Case Requirements ..... 499
4 Sequence Diagram ..... 500
5 References ..... 500
Use Case 11: Industrial AI Quality Inspection Based on IndustryGPT Multimodal Industrial Large Model ..... 502
1 Use Case Summary Table ..... 502
2 Use Case Description ..... 503
3 Use Case Requirements ..... 505
4 Sequence Diagram ..... 505
5 References ..... 505
Use Case 12: AI-Based System for Assessing Health and Moisture Content of Concrete ..... 506
1 Use Case Summary Table ..... 506
2 Use Case Description ..... 507
3 Use Case Requirements ..... 508
4 Sequence Diagram ..... 509
5 References ..... 509

Use Case 13: AI Control System for Fusion Energy ..... 510

1 Use Case Summary Table ..... 510
2 Use Case Description ..... 511
3 Use Case Requirements ..... 512
4 Sequence Diagram ..... 513
5 References ..... 514

4.6. Category 6: Finance ..... 515

Use Case 1: MOF IntelliServe – AI-Powered Conversational & Document Intelligence for Financial Services ..... 515

1 Use Case Summary Table ..... 515
2 Use Case Description ..... 518
3 Use Case Requirements ..... 522
4 Sequence Diagram ..... 523
5 References ..... 524

Use Case 2: AI-Powered Price Negotiation Solution (Nego-Wiz) ..... 525

1 Use Case Summary Table ..... 525
2 Use Case Description ..... 526
3 Use Case Requirements ..... 529
4 Sequence Diagram ..... 530
5 References ..... 530

Use Case 3: AIZEN – CreditConnect-EV ..... 531

1 Use Case Summary Table ..... 531
2 Use Case Description ..... 532
3 Use Case Requirements ..... 534
4 Sequence Diagram ..... 534
5 References ..... 536

Use Case 4: AI-Powered Payroll Solution (AI Payroll Guard) ..... 537

1 Use Case Summary Table ..... 537
2 Use Case Description ..... 537
3 Use Case Requirements ..... 540
4 Sequence Diagram ..... 541
5 References ..... 541

Use Case 5: AI-Driven Financial Inclusion with Enhanced Consumer Protection via Mobile Networks in Zambia ..... 542

1 Use Case Summary Table ..... 542
2 Use Case Description ..... 543
3 Use Case Requirements ..... 544
4 Sequence Diagram ..... 545
5 References ..... 545

Use Case 6: AI for Fraud Detection in NGO Operations: Bringing Trust with Data-Driven Governance ..... 546

1 Use Case Summary Table ..... 546
2 Use Case Description ..... 547
3 Use Case Requirements ..... 551
4 Sequence Diagram ..... 553
5 References ..... 553

Use Case 7: Strengthening Financial Integrity and International Trust: UAE's FATCA/CRS Framework ..... 554

1 Use Case Summary Table ..... 554
2 Use Case Description ..... 555
3 Use Case Requirements ..... 558
4 Sequence Diagram ..... 559
5 References ..... 559

Use Case 8 Sustainable Financing Partnership ..... 560

1 Use Case Summary Table ..... 560
2 Use Case Description ..... 561
3 Use Case Requirements ..... 562
4 Sequence Diagram ..... 562
5 References ..... 562

Use Case 9: Receipt Guardian ..... 563

1 Use Case Summary Table ..... 563
2 Use Case Description ..... 564
3 Use Case Requirements ..... 565
4 Sequence Diagram ..... 565
5 References ..... 567

Use Case 10: Dynamic Threshold Adjustment for Credit Scoring: An Adaptive Optimization Approach for Subprime Lending ..... 568

1 Use Case Summary Table ..... 568
2 Use Case Description ..... 568
3 Use Case Requirements ..... 571
4 Sequence Diagram ..... 572
5 References ..... 572

Use Case 11: InsurVision - Intelligent Insurance Claim Analysis System ..... 573

1 Use Case Summary Table ..... 573
2 Use Case Description ..... 573
3 Use Case Requirements ..... 575
4 Sequence Diagram ..... 575
5 References ..... 575

Use Case 12: AI to lift people out of poverty ..... 577

1 Use Case Summary Table ..... 577
2 Use Case Description ..... 579
3 Use Case Requirements ..... 581
4 Sequence Diagram ..... 583
5 References ..... 583

4.7. Category 7: Education ..... 585

Use case 1: AI-Powered Math Teaching Assistant - Process-level Feedback & Assessment ..... 585

1 Use Case Summary Table ..... 585
2 Use Case Description ..... 586
3 Use Case Requirements ..... 587
4 Sequence Diagram ..... 588
5 References ..... 588

Use case 2: QANDA – An AI Tutor for everyone, advancing equity in education ..... 589

1 Use Case Summary Table ..... 589
2 Use Case Description ..... 590
3 Use Case Requirements ..... 591
4 Sequence Diagram ..... 591
5 Reference ..... 591

Use case 3: AI Helpy for AI Textbook ..... 592

1 Use Case Summary Table ..... 592
2 Use Case Description ..... 592
3 Use Case Requirements ..... 594
4 Sequence Diagram ..... 594
5 References ..... 594

Use case -4: AI-Powered Offline Learning for Higher Education in Crisis Zones ..... 595

1 Use Case Summary Table ..... 595
2 Use Case Description ..... 596
3 Use Case Requirements ..... 597
4 Sequence Diagram ..... 597
5 References ..... 598

Use case -5: AI – Driven Virtual Instructor for 3D Virtual Electronics Lab ..... 599

1 Use Case Summary Table ..... 599
2 Use Case Description ..... 599
3 Use Case Requirements ..... 601
4 Sequence Diagram ..... 603
5 References ..... 603

Use case 6: AI-Powered Personalized Learning for Neurodivergent Students ..... 604

1 Use Case Summary Table ..... 604
2 Use Case Description ..... 605
3 Use Case Requirements ..... 607
4 Sequence Diagram ..... 608
5 References ..... 608

Use case 7: UN Women AI School ..... 609

1 Use Case Summary Table ..... 609
2 Use Case Description ..... 610
3 Use Case Requirements ..... 612
4 Sequence Diagram ..... 612
5 References ..... 612

Use case 8: Little Signs ..... 613

1 Use Case Summary Table ..... 613
2 Use Case Description ..... 614

3 Use Case Requirements ..... 616
4 Sequence Diagram ..... 617
5 References ..... 617
Use case -9: Using AI in Student Creative Art Translations and STEM Workshop ..... 618
1 Use Case Summary Table ..... 618
2 Use Case Description ..... 619
3 Use Case Requirements ..... 620
4 Sequence Diagram ..... 621
5 References ..... 621
Use Case 10: AI-Driven Personalization for Sustainable Development: Transforming Education, Wealth, Nutrition, and Health ..... 623
1 Use Case Summary Table ..... 623
2 Use Case Description ..... 624
3 Use Case Requirements ..... 625
4 Sequence Diagram ..... 626
5 References ..... 626
Use Case 11: Dubai AI & KHDA Smart School Finder ..... 628
1 Use Case Summary Table ..... 628
2 Use Case Description ..... 628
4 Sequence Diagram ..... 630
5 References ..... 630
4.8. Category 8: Smart home/cities ..... 631
Use Case 1: Culturalink: AI-Driven World Cultural Linking and Interactions Engine ..... 631
1 Use Case Summary Table ..... 631
2 Use Case Description ..... 632
3 Use Case Requirements ..... 635
4 Sequence Diagram ..... 636
5 References ..... 636
Use Case 2: Yuanjing Integrated Urban Management Platform ..... 638
1 Use Case Summary Table ..... 638
2 Use Case Description ..... 639
3 Use Case Requirements ..... 641
4 Sequence Diagram ..... 642

5 Reference ..... 642

Use Case 3: AI-Enhanced Detection and Prediction at Saruq Al-Hadid Archaeological Site, Dubai, UAE ..... 643

1 Use Case Summary Table ..... 643
2 Use Case Description ..... 646
3 Use Case Requirements ..... 648
4 Sequence Diagram ..... 650
5 References ..... 650

Use Case 4: Artificial Intelligence + Robots Reconstruct Parking Space and Solve Urban Parking Problems ..... 651

1 Use Case Summary Table ..... 651
2 Use Case Description ..... 651
3 Use Case Requirements ..... 653
4 Sequence Diagram ..... 653
5 Future Work ..... 654
6 References ..... 654

Use Case 5: AI-Powered Video Ringback Tone ..... 655

1 Use Case Summary Table ..... 655
2 Use Case Description ..... 656
3 Use Case Requirements ..... 657
4 Sequence Diagram ..... 658
5 References ..... 659

Use Case 6: YI JIANG JIE: Cultural and Creative Explanation of Big-Scale Model ..... 660

1 Use Case Summary Table ..... 660
2 Use Case Description ..... 661
3 Use Case Requirements ..... 664
4 Sequence Diagram ..... 665
5 References ..... 665

Use Case 7: OneAstronomy[1]: Harnessing Multimodal Astronomical Foundation Models to Pave a New Path for Humanity's Sustainable Development ..... 666

1 Use Case Summary Table ..... 666
2 Use Case Description ..... 667
3 Use Case Requirements ..... 671
4 Sequence Diagram ..... 671

5 References ..... 672

Use Case 8: AI-Based Signal Detection System for Residential TV Dish Alignment ..... 674

1 Use Case Summary Table ..... 674
2 Use Case Description ..... 674
3 Use Case Requirements ..... 676
4 Sequence Diagram ..... 677
5 References ..... 677

Use Case 9: Smart Housing Experience ..... 678

1 Use case Summary Table ..... 678
2 Use Case Description ..... 679
3 Use Case Requirements ..... 681
4 Sequence Diagram ..... 682
5 References ..... 682

Use Case 10: AI-Powered Network Quality Prediction for Seamless Online Event Planning ..... 683

1 Use Case Summary Table ..... 683
2 Use Case Description ..... 684
3 Use Case Requirements ..... 686
4 Sequence Diagram ..... 687
5 References ..... 687

Use Case 11: AI-Powered Space Optimization for Low-Income Housing ..... 688

1 Use Case Summary Table ..... 688
2 Use Case Description ..... 689
3 Use Case Requirements ..... 690
4 Sequence Diagram ..... 691
5 References ..... 691

4.9. Category 9: Accessibility ..... 692

Use Case 1: Bridging the Digital Gap: AI Sign Language Interpretation Solution ..... 692

1 Use Case Summary Table ..... 692
2 Use Case Description ..... 693
3 Use Case Requirements ..... 695
4 Sequence Diagram ..... 696
5 References ..... 696

Use Case 2: AI-Driven Drowsiness Detection: Enhancing Road Safety through Inclusive and Real-Time Monitoring ..... 697

1 Use Case Summary Table ..... 697
2 Use Case Description ..... 698
3 Use Case Requirements ..... 700
4 Sequence Diagram ..... 701
5 References ..... 702

Use Case 3: From Silence to Speech: A Novel AI System for Speech Development for the Deaf Through Lip-Reading ..... 704

1 Use Case Summary Table ..... 704
2 Use Case Description ..... 705
3 Use Case Requirements ..... 707
4 Sequence Diagram ..... 707
5 References ..... 708

Use Case 4: On-Device Intelligence Empowers People with Disabilities and Smart Interaction Scenarios ..... 709

1 Use Case Summary Table ..... 709
2 Use Case Description ..... 710
3 Use Case Requirements ..... 713
4 Sequence Diagram ..... 714
5 References ..... 714

Use Case 5: AI-Powered Behavior Recognition for Developmental Disability ..... 716

1 Use Case Summary Table ..... 716
2 Use Case Description ..... 717
3 Use Case Requirements ..... 718
4 Sequence Diagram ..... 719
5 References ..... 719

Use Case 6: Fine-Tuning ASR Language Models for Underrepresented Low-Resource Languages ..... 720

1 Use Case Summary Table ..... 720
2 Use Case Description ..... 721
3 Use Case Requirements ..... 724
4 Sequence Diagram ..... 724
5 References ..... 725

Use Case 7: Blind Metaverse ..... 726

1 Use case summary table ..... 726
2 Use Case Description ..... 726
3 Use Case Requirements ..... 728
4 Sequence Diagram ..... 729
5 Reference ..... 729

Use Case 8: AI-Powered Indian Sign Language (ISL) Detection and Translation ..... 730

1 Use Case Summary Table ..... 730
2 Use Case Description ..... 732
3 Use Case Requirements ..... 733
4 Sequence Diagram ..... 734
5 References ..... 735

Use Case 9: Intelligent and Automatic System Interface Control System for Users with Disabilities ..... 736

1 Use Case Summary Table ..... 736
2 Use Case Description ..... 737
3 Use Case Requirements ..... 738
4 Sequence Diagram ..... 739
5 Refereneces ..... 739

Use Case 10: Trustworthy Artificial Intelligence for Autonomous Wheelchairs ..... 740

1 Use Case Summary Table ..... 740
2 Use Case Description ..... 741
3 Use Case Requirements ..... 745
4 Sequence Diagram ..... 745
5 Refereneces ..... 746

4.10. Category 10: Intelligent Transport ..... 747

Use Case 1: Tanbeeh: Revolutionizing Safety and Efficiency with AI ..... 747

1 Use Case Summary Table ..... 747
2 Use Case Description ..... 748
3 Use Case Requirements ..... 753
4 Sequence Diagram ..... 754
5 References ..... 754

Use Case 2: Apollo Go: Pioneering Inclusive Mobility for the Visually Impaired ..... 755

1 Use Case Summary Table ..... 755
2 Use Case Description ..... 756
3 Use Case Requirements ..... 758
4 Sequence Diagram ..... 758
5 References ..... 758

Use Case 3: CargoWaves: The AI-Driven Inland Transport Platform Designed to Transform Logistics and Transport Industry ..... 759

1 Use Case Summary Table ..... 759
2 Use Case Description ..... 760
3 Use Case Requirements ..... 763
4 Sequence Diagram ..... 764
5 References ..... 766

Use Case 4: Low-Cost Decision Making for Group of Vehicles for Safe Mobility in V2X ..... 767

1 Use Case Summary Table ..... 767
2 Use Case Description ..... 768
3 Use Case Requirements ..... 772
4 Sequence Diagram ..... 773
5 References ..... 774

Use Case 5: Intelligent In-Vehicle Scenario Orchestration Agent System ..... 775

1 Use Case Summary Table ..... 775
2 Use Case Description ..... 775
3 Use Case Requirements ..... 780
4 Sequence Diagram ..... 781
5 References ..... 781

Use Case 6: Enhancing Road Coordination Through Vehicle-to-Vehicle (V2V) Communication ..... 782

1 Use Case Summary Table ..... 782
2 Use Case Description ..... 783
3 Use Case Requirements ..... 786
4 Sequence Diagram ..... 787
5 References ..... 787

Use Case 7: AI-Driven 6G-Enabled V2X Communications for Smart and Sustainable Mobility ..... 788

1 Use Case Summary Table ..... 788

2 Use Case Description ..... 789
3 Use Case Requirements ..... 792
4 Sequence Diagram ..... 792
5 References ..... 793

Use Case 8: AI for Intrusion in Vehicle-to-Vehicle (V2V) Communication ..... 794

1 Use Case Summary Table ..... 794
2 Use Case Description ..... 795
3 Use Case Requirements ..... 797
4 Sequence Diagram ..... 798
5 References ..... 798

Use Case 9: Environment-Based Endurance Testing in Autonomous Vehicles ..... 799

1 Use Case Summary Table ..... 799
2 Use Case Description ..... 800
3 Use Case Requirements ..... 808
4 Sequence Diagram ..... 809
5 References ..... 809

4.11. Category 11: Smart Agriculture ..... 811

Use Case 1: AI-Enabled Soil Analysis and Weather Station for Local Farmers ..... 811

1 Use Case Summary Table ..... 811
2 Use Case Description ..... 812
3 Use Case Requirements ..... 813
4 Sequence Diagram ..... 814
5 References ..... 815

Use Case 2: IoT-Based Solar-Powered Automatic Basin Irrigation System ..... 816

1 Use Case Summary Table ..... 816
2 Use Case Description ..... 817
3 Use Case Requirements ..... 818
4 Sequence Diagram ..... 819
5 References ..... 819

Use Case 3: AI-Powered Agricultural Contact Center for Farmers ..... 820

1 Use Case Summary Table ..... 820
2 Use Case Description ..... 821
3 Use Case Requirements ..... 822

4 Sequence Diagram ..... 823
5 References ..... 823
Use Case 4: Divine Farmer Agricultural AI Model ..... 825
1 Use Case Summary Table ..... 825
2 Use Case Description ..... 826
3 Use Case Requirements ..... 827
4 Sequence Diagram ..... 828
5 References ..... 828
Use Case 5: Edge AI and IoT-Powered System for Early Tomato Disease Detection and Smart Control ..... 829
1 Use Case Summary Table ..... 829
2 Use Case Description ..... 829
3 Use Case Requirements ..... 831
4 Sequence Diagram ..... 831
5 References ..... 831
Use Case 6: Intelligent Soil Analyzer and Crop Fertilizers Advisory System ..... 833
1 Use case summary Table ..... 833
2 Use case description ..... 833
3 Use case requirements ..... 835
4 Sequence diagram ..... 835
5 References ..... 835

# 1. Introduction

The AI technology is driving a technological revolution, impacting the global economy and society. A wide range of sectors, including healthcare, financial services, manufacturing, agriculture, education, intelligent transportation, 5G networks, and climate change are achieving tangible benefits from integration of AI techniques. These benefits include user friendly applications and services, better usability and performance, availability and agility of services, in addition to better manageability. However, alongside these benefits, experts caution about challenges such as the widening digital divide and the perpetuation of social and algorithmic biases.

To tackle these issues, ITU launched AI for Good Innovate for Impact project initiative under the AI for Good platform. This initiative facilitates sharing of use cases, impactful AI applications and global lessons and expertise from around the world. In 2024, our initiative received positive feedback from both scholars and project contributors, highlighting the value of the selected AI use cases in addressing real-world problems using AI.

Building on this success, we continued the initiative in 2025 by reissuing the **call for AI use cases** and the **call for AI scholars**. In addition, to enhance the scientific rigor of the evaluation process and provide more structured support to participating scholars, we established a new **Technical Advisory Committee (TAC)** composed of experts from various domains. The calls can be accessed on the AI for Good website here.

# 2. Project introduction and selection process

## 2.1 AI use cases

### 2.1.1 Use case selection

Within the deadline for the call for use cases, we received 234 submissions from 32 countries. For evaluation of the use cases, a set of criteria was established encompassing clarity and completeness, justification, alignment with AI for Good objectives, and evaluation against specific metrics, each with an assigned percentage. Considering our timeline constraints, we decided to release an interim collection of use cases in July 2025, featuring an initial 160 use cases. Subsequently, we will continue the engagement with use case authors and scholars to enhance and refine the collection, with the final version scheduled for release in September 2025.

### 2.1.2 Geographical distribution

The 160 use cases came from 326 authors from 28 countries of six regions, including Australia, Canada, China, Cyprus, Côte d'Ivoire, Ethiopia, Georgia, Ghana, India, Iraq, Italy, Korea, Lebanon, Luxembourg, Malaysia, Nigeria, Palestine, Russia, Saudi Arabia, Switzerland, Tanzania, Thailand, Uganda, United Arab Emirates, United Kingdom, United States, Zambia, and Zimbabwe

This geographical diversity underscores the global momentum behind AI adoption. Asia and Pacific region lead the charge, powered by established tech giants and rapidly rising innovation hubs across developing countries. Africa's strong representation reflects its successful embrace of digital technologies, demonstrating how AI can help leapfrog traditional development pathways and accelerate digital transformation. In Arab state region, active participation highlights strategic investments in AI as part of broader efforts to modernize economies and reduce reliance on natural resources. Europe and CIS region contributes through well-developed AI ecosystems that integrate cutting-edge research with robust regulatory frameworks, while Americas remains a driving force in foundational innovation and large-scale enterprise applications.

From 326 authors who submitted use cases, we extended invitations to 45 authors to join us at the Summit and showcase their work.

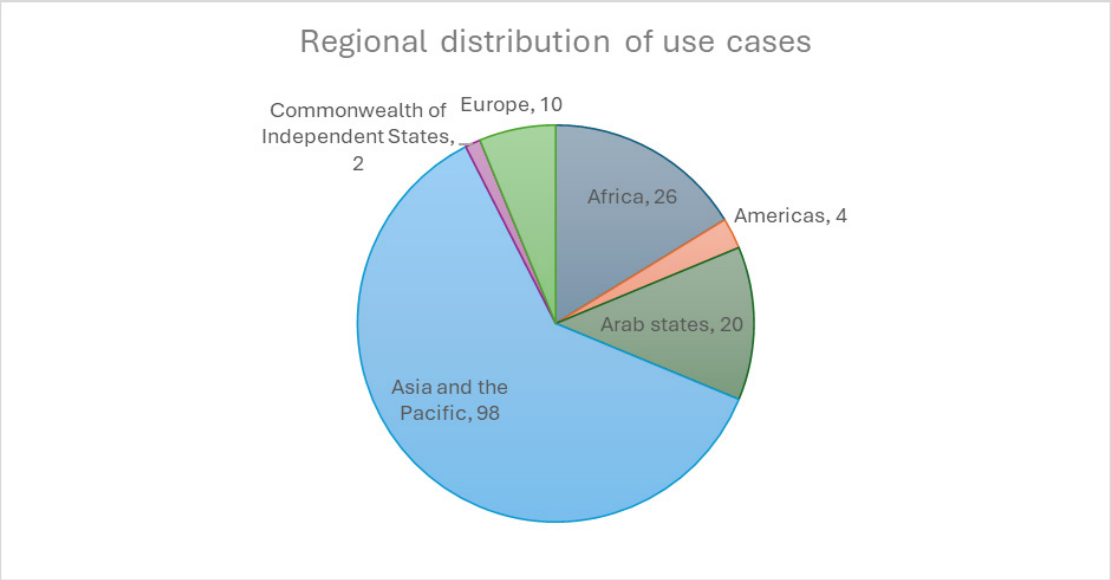



Figure 1: Regional Distribution of Use Cases

### 2.1.3 Distribution of use cases by domains

The AI use cases have been broadly categorized into **11 application domains**, offering a clear lens through which to assess global priorities and technological readiness. The data reveals a distinct pattern of emphasis that aligns with both urgent societal needs and the maturity of AI adoption across sectors.

**Healthcare** stands out as the leading domain, featuring the highest number of cases. This highlights AI's critical role in addressing health challenges through medical diagnostics, treatment optimization, and pandemic preparedness–issues that have grown increasingly prominent in recent years.

**Climate and natural disaster management** ranks as the second most represented area, reflecting the rising importance of AI in climate modelling, environmental monitoring, disaster response, and sustainability efforts. This trend underscores global concern over climate change and the demand for intelligent systems that can support mitigation and adaptation strategies.

The **telecommunications** sector also demonstrates significant activity, particularly in the context of 5G deployment, where AI is enhancing infrastructure, network optimization, and smart connectivity. **Productivity enhancement** features prominently, indicating widespread efforts to use AI for process automation, workflow optimization, and augmenting human performance across various industries.

In **manufacturing**, AI is driving industrial transformation through smart factories, predictive maintenance, and intelligent automation. The **financial services** sector showcases AI applications in fraud detection, credit scoring, risk analysis, and inclusive finance.

**Education** and **smart cities and homes** represent growing areas of interest, demonstrating how AI is reshaping both learning environments and urban living. **Accessibility solutions** illustrate AI's potential to create inclusive technologies for people with disabilities, while **transportation** applications focus on autonomous driving, traffic optimization, and logistics.

Finally, although **smart agriculture** appears less frequently, its impact is significant—particularly in developing regions where AI supports food security and sustainable farming.

This distribution suggests a thoughtful and strategic approach to AI deployment, prioritizing immediate human needs like healthcare and climate resilience while laying the groundwork for long-term innovation, infrastructure development, and inclusive growth across all sectors.

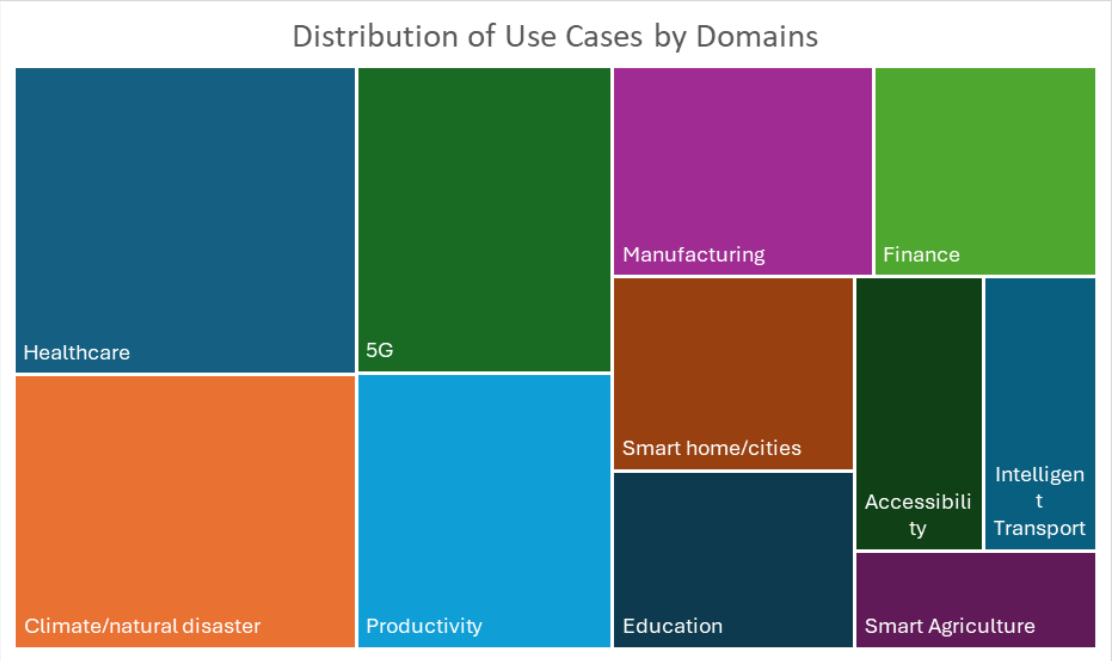



Figure 2: Distribution of use cases by domains

### 2.1.4 Entities who submitted AI use cases

From both the perspective of entities engagement and Distribution of use cases by entity categories, a diverse and synergistic AI ecosystem has taken shape. Industry leads in driving implementation, showcasing strong engagement and a high output of practical applications. Academia remains a cornerstone of innovation, contributing significantly through research and talent development.

Government bodies demonstrate proactive involvement, using AI to advance public services and policy frameworks. NGOs and UN organizations offer high-impact contributions focused on ethics, inclusion, and global priorities.

Together, these entities form a complementary structure where each plays a distinct yet interconnected role, fostering responsible and sustainable AI development through a wide range of real-world use cases.

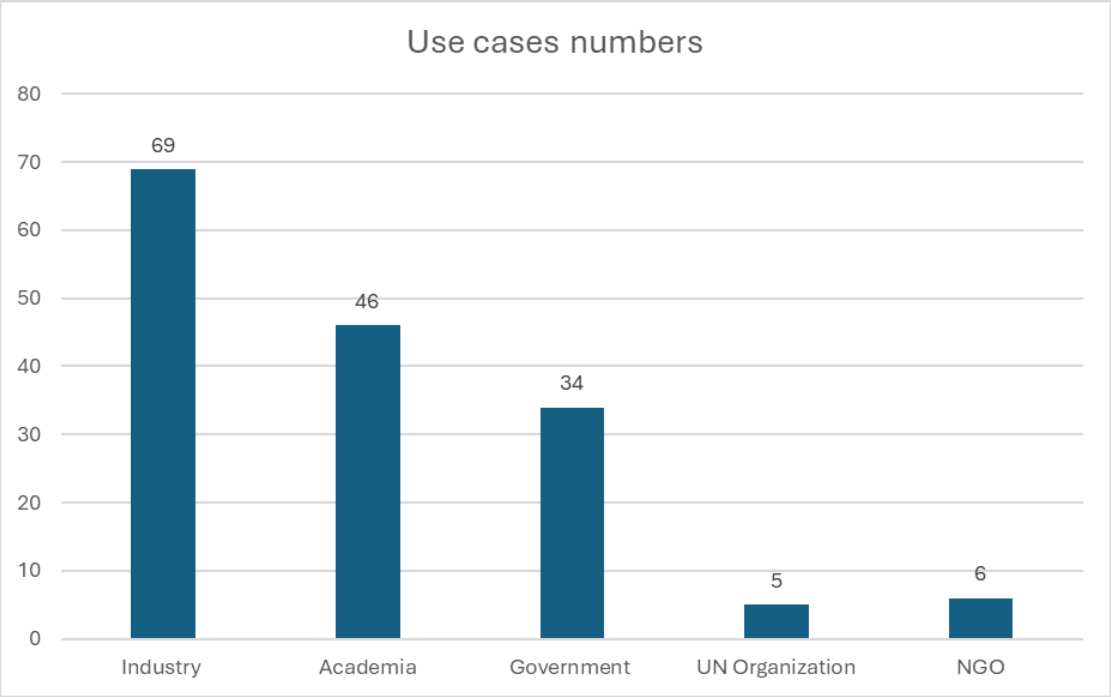



Figure 3: Distribution of use cases by entity categories

## 2.2 Technical Advisory Committee formation and Selection

To strengthen the scientific rigor and real-world relevance of project evaluation, a **Technical Advisory Committee (TAC)** was established in December 2024. The committee brings together experts from diverse disciplines, including AI research, engineering, ethics, social sciences, and sustainable development. Within the call for participation, we received expressions of interest and nominations from a wide range of professionals.

The **selection criteria** included domain expertise, experience in applying AI for social impact, regional and gender balance, and demonstrated commitment to interdisciplinary collaboration. After careful review, we selected **32 TAC members** (see Table <1>) representing academic institutions, think tanks, industry, and international initiatives. These members were invited to support the *Innovate for Impact 2025* cycle.

The TAC plays a central role in reviewing submitted use cases, advising on pilot project selection, and contributing to the strategic direction of the initiative. It also helps identify emerging trends, ensure methodological soundness, and promote inclusive and responsible AI development. Some of TAC members would attend the AI for Good Global Summit 2025 and Innovate for Impact workshop.

Table 1

| | Name | Country |
|---|---|---|
| 1 | Paula Alexandra Palade | United Kingdom |
| 2 | Fikret Sivrikaya | Germany |
| 3 | Artem Volkov | Russian Federation |

## Table 1 (continued)

| | Name | Country |
|---|---|---|
| 4 | Ammar Muthanna | Russian Federation |
| 5 | Halima Mohamed Ismaeel | Bahrain |
| 6 | Dennis Agyemanh Nana Gookyi | Ghana |
| 7 | Aduwati binti Sali | Malaysia |
| 8 | Lina Bariah | United Arab Emirates |
| 9 | Kiran Raj Pandey | Nepal |
| 10 | Katarzyna Wac | Switzerland |
| 11 | Sonali Garg | India |
| 12 | Abhishek Girish Dandekar | Germany |
| 13 | Ashish Sanjay Sharma | Sweden |
| 14 | SARA EL-ATEIF | Morocco |
| 15 | Shweta Khushu | Canada |
| 16 | Antony Franklin Antony Pudhumai Anandam | India |
| 17 | LIU Ao | China |
| 18 | QIU Yue | China |
| 19 | Yu Rong | China |
| 20 | Pei-Yun Sabrina Hsueh | United States |
| 21 | Dr. Asrat Mulatu Beyene | Ethiopia |
| 22 | Moon Jeong Choi moonchoi | Korea |
| 23 | Sameer Kulkarni | India |
| 24 | Yi ZENG | China |
| 25 | Ameer Shohail Liakath Hussain | Saudi Arabia |
| 26 | Prof. James Agajo | Nigeria |
| 27 | Michael Mollel | Tanzania |
| 28 | Ameni Khachlouf | Tunisia |
| 29 | James Boogere | Uganda |
| 30 | Daniel Rezende | Brazil |
| 31 | Marcelo Mendoza | Chile |
| 32 | Patrick McSharry | United Kingdom |

## 2.3 Scholars Selection

Within the deadline for the Call for AI4G Scholars, we received 171 scholar applications from 37 countries. We aimed for identifying top candidates who embody a diverse range of qualities and commitments. The selection criteria were academic qualifications, interests and dedication to AI technology, collaboration and partnership, communication skills, and a recommendation letter. We selected 17 scholars from academia, universities, and AI companies (see Table <2>) who were invited to participate in the AI for Good Global Summit 8-11 July 2025. Scholars worked closely with use case authors to understand the technology, the use case and the domain. They received feedback and mentoring to derive valuable insights in collaboration with the experts. We will award selected scholars for their outstanding efforts including Elzahraa Ibrahim Mohamed Abdelhafiz, ANNU, Linbi DENG, Yi CHEN, Syed Azim, Innocent Nzimenyera and Shadia Y M Baroud.

### Table 2: Seventeen AI scholars were selected

| | Name | Country |
|---|---|---|
| 1 | Bereket Teklew Kibret | Ethiopia |
| 2 | Ndabuye Sengayo Gideon | Tanzania |
| 3 | Syed Azim | India |
| 4 | Elzahraa Ibrahim Mohamed Abdelhafi | Egypt |
| 5 | Linbi Deng | China |
| 6 | Ayman Babikir Elzaki Mohamed | UAE |
| 7 | ANNU | India |
| 8 | Yi Chen | China |
| 9 | Dileesh chandra bikkasani | USA |
| 10 | Innocent Nzimenyera | Rwanda |
| 11 | Meghana Kotcherlakota | USA |
| 12 | ADEGOKE ADEDOLAPO ISRAEL | Nigeria |
| 13 | Nchimunya Kabalo | Zambia |
| 14 | Naome Etori | Kenya |
| 15 | Theofrida Maginga | Tanzania |
| 16 | Shadia Y. M. Baroud | Palestine[1] |
| 17 | Aizhan Tursunbayeva | Italy |

[1] RESOLUTION 99 (REV. DUBAI, 2018) Status of Palestine in ITU

# 3. Learnings and conclusion

## 3.1. Cross-Sector Insights from Global Use Cases

This report presents a synthesis of 160 real-world AI use cases across eleven critical domains, offering a comprehensive view of how AI is being applied to address societal challenges, drive innovation, and enhance efficiency. It highlights both the transformative impact of AI across sectors and the persistent gaps—technical, governance and infrastructural—that must be bridged to achieve scalable, inclusive, and sustainable deployment worldwide.

1. An analysis of 27 healthcare AI use cases highlights three key barriers to large-scale adoption: lack of data standardization, limited real-world deployment of edge AI, and insufficient explainability in multimodal models. Despite these challenges, the field is advancing rapidly with trends such as specialized large models, generative AI for medical data, federated learning for privacy-preserving training, and lightweight AI for mobile health. Some models now match or surpass expert performance, especially in multi-cancer detection, signalling a shift toward more intelligent, inclusive, and deployable AI-driven healthcare solutions.
2. Over 23 AI use cases in climate and disaster response reveal major challenges, including inconsistent data standards, regional data scarcity, and limited global scalability. Nonetheless, the project has helped collect valuable data from underrepresented regions, enhancing model generalization and supporting more locally tailored, effective solutions. Through cross-border collaboration and data sharing, AI is driving a more inclusive and resilient global response to climate risks.
3. AI use cases in the 5G domain face three key challenges: infrastructure integration across heterogeneous systems, real-time performance limitations, and compliance and security risks in cross-domain applications. Still, a dual transformation is underway. "Network for AI" leverages 5G/6G infrastructure—cloud-edge collaboration, network slicing, and programmable interfaces—to support scalable AI workloads. In parallel, "AI for Network" enables intelligent operations like predictive maintenance and dynamic resource management, reshaping telecom infrastructure into an adaptive, self-optimizing platform.
4. More than 20 AI productivity use cases illustrate a shift from experimental development to real-world deployment. These applications—from smart HR systems to multilingual translation tools—feature mature architectures, scalable integration, and growing adoption. While challenges around data quality and localization persist, they are being addressed through human-in-the-loop methods and domain adaptation. Trends like API integration and open-source adoption signal that AI is now a key driver of cross-sector productivity.
5. Current AI use cases in manufacturing highlight the convergence of AI, IoT, and robotics to enable smart inspection, autonomous operations, and predictive maintenance. However, challenges remain in communication reliability, data availability, safety standards, and trust in automation. Addressing these issues is essential to realize intelligent, adaptive industrial ecosystems at scale.
6. Financial AI use cases raise concerns about the balance between personalization and privacy. Intelligent services such as robo-advisors and voice-activated banking depend on sensitive user data yet often lack transparency and adequate control mechanisms. Best practices—including on-device AI, differential privacy, federated learning, and explainable AI—are emerging to address these gaps and build secure, ethical, and user-centered financial services.
7. Many AI education solutions remain in pilot phases, constrained by scalability issues, infrastructure gaps, and misalignment with local curricula. While tools like AI tutors and

adaptive platforms show promise, structured pilots, phased implementation, and teacher feedback are key to building trust and enabling sustainable adoption.

8. In smart city and home environments, user awareness and usability are critical barriers. Many users struggle to effectively interact with AI systems due to unfamiliarity or lack of confidence. Without user-centric design and user awareness and capacity building efforts, even advanced features risk underutilization. Enhancing transparency, localization, and user support is essential to maximize real-world impact.
9. AI accessibility use cases—ranging from mobile sign language translation to multimodal assistive devices—demonstrate strong potential for inclusion. However, challenges remain in hardware compatibility, regional data diversity, and user trust. Addressing these requires lightweight model optimization, localized data, privacy safeguards, and sustained deployment to ensure inclusive adoption.
10. AI in intelligent transport faces key gaps in large-scale deployment, real-time data integration, and environmental impact. Yet, drone and UAV use cases are emerging as agile, low-emission alternatives, and AI-powered energy solutions—such as traffic optimization and smart EV charging—are proving vital for reducing fuel consumption and supporting sustainable mobility.
11. Many smart agriculture AI use cases focus on rural settings, especially in Africa and Southeast Asia. These applications tackle productivity and digital access challenges through low-cost, offline-capable tools such as voice-enabled advisory platforms and edge-based disease detection. In doing so, they not only support local farmers but also enrich global agricultural AI systems with diverse, representative data.

## 3.2. Key Insights from 2025 Innovate for Impact analysis

**1. Regional requirements and locally relevant data:** The review of 160 AI use cases across eleven domains reveals valuable lessons on how AI can move from concept to real-world impact. A key takeaway is that successful implementation depends not only on technical innovation, but on understanding regional requirements and locally relevant data. Many impactful projects emerged from tailoring AI tools to the unique challenges, infrastructure, and linguistic contexts of specific communities.

**2. Capacity building:** Another consistent pattern is the importance of capacity building, especially through user training and knowledge sharing. Use cases demonstrate how providing practical examples, reusable templates, and locally adapted tools can significantly improve adoption and long-term sustainability. Mentoring and guidance from experts played an important part in producing quality results.

**3. Supportive ecosystem**: In addition, the role of a supportive ecosystem—including access to mentoring, cloud infrastructure, and financial assistance such as scholarships or pilot funding—proves critical to scaling AI solutions. Many initiatives benefited from early-stage guidance and resource access that allowed them to move beyond experimentation toward implementation. Providing compute resources to developing country researchers is important to bridge the gap in ability to solve local problems using AI.

Together, these learnings highlight that AI for Good is not simply about technology– it gives us an opportunity for reaching out to regional innovators, creating the right conditions for inclusive, effective, and context-sensitive deployment that delivers measurable benefits to communities around the world.

## 3.3. Future work

The trajectory of work over the coming years reflects a clear evolution from ideation to implementation and ultimately to scalable integration:

- **2024 - Concept based use cases:** in 2024, we focused on collecting and analysing concept-based use cases. These examples offered insights into potential AI applications across diverse domains, serving as a foundation for understanding opportunities and identifying key challenges.
- **2025 - Practical pilots and real-world implementations:** in 2025, we emphasize practical pilots and real-world implementations, testing selected use cases in live environments. The goal is to validate functionality, assess impact, and gather user feedback to refine solutions and demonstrate feasibility at scale.
- **2026 - Interoperability and scaling:** in 2026 we aim to support the development of open, interoperable, and widely applicable AI approaches. By encouraging knowledge sharing and common frameworks, this phase will help translate individual successes into adaptable models that can be widely applied across regions and sectors.

Considering this progress, we call for engagement from the industry, academia, member States and other members for this project. This call will involve collaboration around the next steps of these use cases, including implementation guidelines, deployment stories, opensource respositories and validation toolsets. Expanded support for developing countries, including cloud credits for computing and training of AI/ML models would enable more inclusive development of AI around the world.

# 4. Innovate for impact use cases

## 4.1. Category 1: Healthcare

### Use Case 1: Intelligent Applications Based on Medical Large Models to Enhance People's Healthcare Experience and Health Status

**Organization:** Alipay (Hangzhou) Information Technology Co., Ltd.

**Country:** China

**Contact Person(s):**

He Sheng, shenghe.shh@antgroup.com

Nannan DU, nannan.dnn@antgroup.com

#### 1 Use Case Summary Table

| Item | Details |
|---|---|
| Category | Healthcare |
| Problem Addressed | Patient experience compromised due to the gap between supply and demand for medical resources.<br>Complex and administrative workflow limits doctors' clinical focus.<br>Uneven geographic distribution of medical resources. |
| Key Aspects of Solution | For Residents: This solution introduces an AI health assistant that covers the entire medical journey, allowing everyone to have a smart family doctor.<br>For Doctors: The solution offers doctors an intelligent assistant to aid in clinical care and administrative tasks, becoming a valuable helper.<br>For Underserved Areas: This case establishes intelligent doctor agents that reach more patients online, balance healthcare resource distribution, and enhance accessibility and efficiency of high-quality healthcare services. |

(continued)

| Item | Details |
|---|---|
| Technology Keywords | RAG, high-quality medical datasets, AI agents |
| Data Availability | RJUA-QADatasets: Ant-Renji Urology Specialty QA Dataset [12] |
| Metadata (Type of Data) | structured and unstructured data |
| Model Training and Fine-Tuning | The model is pretrained on high-quality, domain-specific medical data, including both textual and multimodal sources such as visually rich clinical documents.<br>The model is then finetuned using expert annotated QA pairs and multi-turn dialogues which adapted to various downstream medical application scenarios. |
| Testbeds or Pilot Deployments | Zhejiang Provincial Health Commission [13]<br>Hangzhou Municipal Medical Insurance Bureau [14]<br>Shanghai Jiao Tong University School of Medicine, Renji Hospital [15] |
| Code repositories | Code repositories are available [16] [17] |

# 2 Use Case Description

## 2.1 Description

### Challenges

- Patient experience compromised due to the gap between supply and demand for medical resources: As living standards rise, the demand for healthcare services has increased in both quantity and quality, but the growth of medical resources has not fully kept pace with this change.
- Complex and administrative workflow limits doctors' clinical focus: In traditional healthcare sector, doctors have to deal with complex workflows, limiting the time they can dedicate to diagnosis and treatment.
- Uneven geographic distribution of medical resources: There are disparities in the geographic distribution of healthcare resources, affecting the accessibility and convenience of high-quality healthcare for residents in different regions.

These factors combined have impacted on the level of healthcare available to residents across different areas.

### Solutions

To address the above challenges, this case study proposes an intelligent application solution based on Ant Group's large medical model, aiming optimize healthcare resources through technology, promote universal access to smart healthcare services, and advance wellbeing for all.

- For Residents: This solution introduces an AI health assistant that covers the entire medical journey, allowing everyone to have a smart family doctor. For patients with explicit

healthcare needs, AI health assistant can provide comprehensive guidance and support throughout the medical process, including triage, insurance consultation, and report interpretation. For residents with general health needs, AI health assistant can offer daily consultations on exercise, diet, mental health, sleep, and more.

- For Doctors: The solution offers doctors an intelligent assistant to aid in clinical care and administrative tasks. The solution's value lies in optimizing doctors' current "70% administrative work + 30% clinical work" time allocation. AI handles tasks it excels at—such as standardizing information and streamlining processes—freeing doctors to focus their expertise on complex case analysis and doctor-patient communication, thereby release more medical resources.
- For underserved areas: This case collaborates with top-tier medical teams to input the clinical experience of national renowned doctors and departments into the model, creating doctor agents or specialty-specific agents. Ordinary patients can access medical advices without visiting hospitals. For complex cases or when in-person visits are necessary, the intelligent agent can recommend nearby primary healthcare facilities, and local doctors can combine the agent's symptom collection and insights to deliver diagnoses, improving both the efficiency and quality and also mitigating regional disparities.

All the above-mentioned solutions are built on the underlying capabilities of Ant Group's medical large model. We will elaborate the technological innovations within the model in the following section.

## Technological Innovations

- **High-Quality Medical Knowledge Graph and Professional Annotation Process**:
  - ***High-Quality Medical Data***: To tackle the challenges of scarce high-quality medical data, the project collaborated with hundreds of professional medical teams to build a data system containing hundreds of billions of tokens of professional medical texts, hundreds of millions of multimodal images and texts, and tens of millions of high-quality medical knowledge graph entries, covering the full pre-, during-, and post-diagnosis cycle with multimodal data formats (text, image, audio, video).
  - ***Professional Annotation Process***: Pioneered a fully automated annotation system combining large model distillation technology and deep involvement from specialist doctors, improving annotation efficiency by 50%. The RJUA series medical datasets released with Shanghai Renji Hospital cover 85% of disease types, 97% of patient demographics, and 96% of high-frequency lab items, setting an industry benchmark. The project's data annotation and evaluation system also create a feedback loop to continuously enhance data quality.
- Multimodal **Interaction Accuracy Exceeding 90%, Achieving Industry-Leading Medical Expertise**:
  - ***Multimodal Interaction***: Built on Ant Group's self-developed medical large language model and training optimization algorithms, the system supports multimodal data exchange including images (e.g., medical imaging with resolution, timestamp, device info, anonymized patient IDs, annotations), audio (e.g., symptom descriptions, doctor-patient conversations), and video (e.g., rehabilitation training, remote consultations).
  - ***Model Accuracy Over 90%***: Using KG-CoT (Knowledge Graph Chain of Thought) and PRM-MCTS (Process Reward Models and Monte Carlo Tree Sear) technologies, the system achieves real-time optimization of diagnostic pathways and 98% tool invocation accuracy. A Unified-Alignment Post-Training framework deeply integrates reinforcement learning and instruction fine-tuning, achieving 93% diagnostic accuracy and over 95% personalized service satisfaction.

- ***Medical Professionalism***: The model surpasses GPT-4 in Chinese and English medical evaluations, ranking first on PromptCBLUE-A and the Chinese medical benchmark MedBench.

- **Safety and Compliance:**
  - ***Data Privacy Protection***: By combining data perturbation, differential privacy, and multi-role model rewriting, the project ensures fuzzy encryption of patient privacy while retaining critical medical information. Multimodal data is linked via unified patient IDs, session IDs, and timelines under strict privacy protocols, ensuring coherence and contextual awareness.
  - ***Verification and Testing Platform***: The project's verification platform standardizes patient creation for real, repeatable testing scenarios, evaluating AI's abilities in comprehension, response, and healthcare support accuracy. All datasets include detailed metadata, and accuracy, reliability, and usability are tested across mobile, web, and smart device platforms to ensure safety, compliance, and clinical value.
  - ***Cautious Human-AI Collaboration***: Adds prudent guidance layers on top of large model outputs, combining expert and AI collaboration, with safety and control mechanisms in place.

- **Doctor Digital Twin for Efficient Communication**:
  - ***Edge-Cloud Collaborative AI Avatar***: Combining diffusion models with dynamic neural radiance fields, the project supports highly realistic digital humans generated from a single photo, achieving an industry-leading low Frechet distance of 17.06. The self-developed edge-cloud rendering solution cuts concurrent service costs by 90%, enabling the large-scale application of doctor digital humans in health management scenarios.
  - ***Audio Interaction with*** National Renowned Doctors: Trained a medical knowledge-enhanced voice large model achieving 97% speech recognition accuracy. Proprietary one-shot voice cloning technology reduced the cost of doctor voice customization by 95%. The end-to-end native voice model keeps duplex conversation response latency within one second, offering a realistic interaction experience for health management.

Use Case Status: The use case is part of a larger product development

## 2.2 Benefits of the use case

1. Lowering the Barriers to Healthcare Access and Bridging the Digital Divide

- By leveraging large-scale visual language recognition with trillions of parameters and advanced learning from vast medical Q&A datasets, multimodal interactions are made accessible to users of all ages and technical backgrounds. Since its launch in June 2024, the service has reached over 13 million users, delivered more than 78 million service instances, and achieved a 98% user satisfaction rate.
- An open collaboration plan for intelligent agents has been introduced to reduce costs for medical institutions and developers. In terms of current implementation, this case integrates intelligent agents from local health commissions and hospitals, offering medical guidance across over 1,000 healthcare facilities.

2. Optimizing Resource Allocation and Advancing Equitable Healthcare

- The system enables patients to quickly identify the most suitable medical resources based on their symptoms and medical history.

- By collaborating with primary healthcare institutions, it expands access to high-quality medical applications in remote and underserved areas.

Its deployment in China—a vast, healthcare ecosystem— brings an inspirational approach: AI may dissolve spatiotemporal constraints in medical resource distribution, ensuring equitable access to inclusive care for all individuals.

### Partners

The project partners with Zhejiang Provincial Health Commission, Hangzhou Municipal Medical Insurance Bureau, and Shanghai Jiao Tong University School of Medicine Renji Hospital. Collaboration with these parties enhances the application's medical professionalism.

## 2.3 Future Work

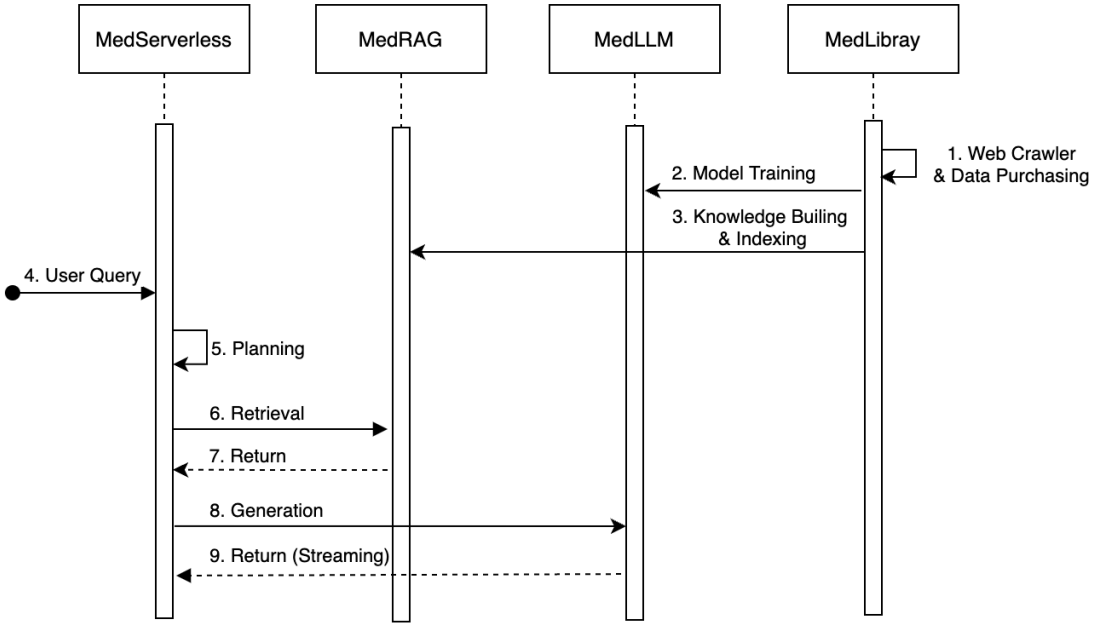
The future work on the intelligent applications use case could expand across several key areas:

- **Standards development related to the use case:** The resources will be used to strengthen collaboration with the International Telecommunication Union (ITU) and promote the development of relevant ITU standards. This will enhance the accessibility and service quality of medical large models, contributing to the achievement of Universal Health Coverage. By leveraging standards related to medical large models, we aim to explore viable pathways for applying AI across various healthcare service fields.
- **Data collection and model development**: These resources will enhance the quality of medical training data, refine and improve the models through continuous fine-tuning and post-training, and strengthen the capabilities of medical knowledge graphs and intelligent agents.

# 3 Use Case Requirements

- **REQ-01:** The medical large models behind the intelligent applications can interact with users through text, images, and voice, instead of other modalities.
- **REQ-02**: The intelligent applications were developed to provide strong language processing and generation capabilities, as well as medical service invocation capabilities, for users in China. It currently supports Chinese and English, and it may not provide strong services for other languages.
- **REQ-03:** It is critical that the usage of the medical large models must comply with all local regulations and guidelines related to digital services, data handling, and language use with China.

## 4 Sequence Diagram




## 5 References

[1] *Towards Automatic Evaluation for LLMs' Clinical Capabilities: Metric, Data, and Algorithm*. 2024. arXiv. https://arxiv.org/abs/2403.16446.

[2] *Learning to Plan for Retrieval-Augmented Large Language Models from Knowledge Graphs*. 2024. arXiv. https://arxiv.org/abs/2406.14282.

[3] *Towards Structured Understanding of Marketer Demands with Analogical Reasoning Augmented LLMs*. 2024. arXiv. https://arxiv.org/abs/2401.04319.

[4] *KnowAgent: Knowledge-Augmented Planning for LLM-Based Agents*. 2024. arXiv. https://arxiv.org/abs/2403.03101.

[5] *RJUA-MedDQA: A Multimodal Benchmark for Medical Document Question Answering and Clinical Reasoning*. 2024. ACM Digital Library. https://dl.acm.org/doi/10.1145/3637528.3671644.

[6] *Effectively PAIRing LLMs with Online Marketing via Progressive Prompting Augmentation*. 2023. arXiv. https://arxiv.org/pdf/2312.05276.

[7] *FoRAG: Factuality-Optimized Retrieval Augmented Generation for Web-Enhanced Long-form Question Answering*. 2024. ACM Digital Library. https://dl.acm.org/doi/10.1145/3637528.3672065.

[8] *EDiT: A Local-SGD-Based Efficient Distributed Training Method for Large Language Models*. 2024. OpenReview. https://openreview.net/forum?id=xtlMtbVfWu.

[9] *Customizable Combination of Parameter-Efficient Modules for Multi-Task Learning*. 2023. arXiv. https://arxiv.org/abs/2312.03248.

[10] *Every FLOP Counts: Scaling a 300B Mixture-of-Experts LING LLM without Premium GPUs*. 2025. arXiv. https://arxiv.org/abs/2503.05139.

[11] Alipay. 2024. "RJUA_Ant_QA." GitHub. https://github.com/alipay/RJU_Ant_QA.

[12] Alipay-Med. 2024. "medDQA_benchmark." GitHub. https://github.com/Alipay-Med/medDQA_benchmark.

[13] Alipay-Med. 2024. "SPs_benchmark." GitHub. https://github.com/Alipay-Med/SPs_benchmark.

[14] Zhejiang Provincial Health Commission. 2023. "Zhejiang Provincial Health Commission Releases Research Results on Medical Document Question Answering." https://wsjkw.zj.gov.cn/art/2023/12/15/art_1229727989_59021387.html.

[15] Hangzhou Municipal Government. 2023. "Hangzhou Accelerates the Development of Medical Artificial Intelligence." https://www.hangzhou.gov.cn/art/2023/2/15/art_812269_59074666.html?eqid=ff95f40c0000f1af00000004647d5097.

[16] WeChat Official Account. 2023. "Alibaba Medical Team Releases Multimodal Medical Question Answering System." https://mp.weixin.qq.com/s/KU0O05UlF6scxlRUd3_Ljw.

# Use Case 2: AI based Virtual Yoga Instructor for Good health

**Organization:** Centre for Development of Advanced Computing, Pune

**Country**: India

**Contact Person(s)**:

Lakshmi Sandip Panat, lakshmip@cdac.in

Swapna Yenishetti, swapnae@cdac.in

Ganesh Karajkhede, ganeshk@cdac.in

## 1 Use Case Summary Table

| Item | Details |
|---|---|
| Category | Healthcare |
| Problem Addressed | Yoga learners and practitioners often face challenges of finding convenient time slots for yoga practice, travel time required to reach the yoga studio, personalized yoga coaching, tracking the improvement in performing the yoga pose, affordability of yoga classes, and customized content for specific health conditions like yoga for diabetes, yoga for heart disease, etc. Also, the cost of performing yoga poses incorrectly can result in negative effects on the body, leading to muscle injury, joint pain, back pain, nerve compression, pressure on internal organs, leading to injury, etc. Yoga learners will benefit |
| Key Aspects of Solution | • Real-time Yoga Pose Correction: Develop an AI algorithm to provide immediate feedback on yoga postures during practice.<br>• Improved Health Outcomes: Enable practitioners to achieve better health benefits by ensuring correct and safe pose execution.<br>• Reduced Reliance on Human Instructors: Minimize the necessity for constant human supervision, making yoga practice more accessible.<br>• Dynamic and Lightweight Processing: Ensure the system can handle the dynamic nature of yoga poses with real-time processing capabilities while remaining efficient and resource-friendly. |
| Technology Keywords | Yoga Asanas, Yoga Pose correction, Convolutional Neural Network (CNN), Movenet, TensorFlow/Keras [5] |
| Data Availability | Yoga data publicly available [1] |
| Metadata (Type of Data) | Images, Landmarks and Video Data |

(continued)

| Item | Details |
|---|---|
| Model Training and Fine-Tuning | A convolutional neural network (CNN) model, MoveNet, was used for processing videos generated while practising yoga poses. This algorithm achieves a good level of accuracy and also provides multi-person support. A custom dataset was prepared consisting of 10 yoga poses: Bhujangasana, Garudasana, Halasana, Natarajasana, Sukhasana, Makarasana, Savasana, Simhasana, and Tadasana. In addition, the Yoga-82 dataset was studied [2]. |
| Testbeds or Pilot Deployments | Active, Demo available on request |
| Code repositories | [1] |

# 2 Use Case Description

## 2.1 Description

Yoga, recognised globally as an art and science for healthy living, has seen a significant surge in popularity, underscored by the United Nations' declaration of June 21st as International Day of Yoga. While its benefits are widely acknowledged, effective practice hinges on expert guidance and consistent adherence. However, numerous challenges hinder widespread access, including the difficulty of finding convenient class times, the need for travel to studios, the lack of personalised coaching, the inability to track individual progress, the high cost of classes, and the scarcity of customised content for specific health conditions like diabetes or heart disease. Critically, incorrect posture during yoga can lead to adverse physical consequences, such as muscle injury, joint pain, back pain, nerve compression, and even internal organ damage. Therefore, there is a clear demand for solutions that enable learners to focus on proper alignment and breathing, ensuring a safe and effective yoga practice [3]. To address these limitations, an innovative technological approach is essential. Our primary objective is to develop an AI-driven algorithm capable of providing real-time yoga pose correction. This algorithm will offer immediate feedback to practitioners, leading to improved health outcomes and significantly reducing the reliance on human instructors. Given the dynamic nature of yoga poses, the system must be designed for real-time processing and maintain a lightweight architecture to ensure seamless performance. Our innovative solution leverages the Convolutional Neural Network (CNN) model, specifically MoveNet, for processing videos of yoga practice. This algorithm is selected for its high accuracy and its ability to support multi-person scenarios. To facilitate its training, a custom dataset was meticulously prepared, encompassing ten fundamental yoga poses: Bhujangasana, Garudasana, Halasana, Natarajasana, Sukhasana, Makarasana, Savasana, Simhasana, and Tadasana. Furthermore, we have incorporated insights from Yoga-82, a comprehensive dataset designed for fine-grained classification of human poses, which offers a three-level hierarchy encompassing body positions, variations, and specific pose names. This technological approach forms the core of our "Virtual Yoga Teacher with AI Model for Yoga Posture correction for Good health," a concept elaborated upon in our publication by Swapna Yenishetti, Ganesh Karajkhede, and Lakshmi Panat. The integration of AI, particularly CNN models, offers profound benefits for both preventive and curative healthcare. As a powerful tool for real-time human pose estimation and tracking, AI significantly enhances the outreach of yoga, transforming it into a non-invasive and sustainable healthcare model accessible to

a broader populace. This technology can democratize healthcare solutions, making them affordable and personalized for citizens across all ages, genders, economic strata, and social backgrounds. By providing accessible and mobile-based AI solutions, public health education in yoga care can be scaled to a population level, thereby alleviating the burden on hospitals and optimizing the utilization of human and budgetary resources. Crucially, AI-powered yoga applications will play a vital role in controlling non-communicable diseases (NCDs) such as diabetes and cardiovascular conditions with unprecedented speed and scale, offering a cost-effective solution particularly beneficial for developing countries. The availability of data-driven health monitoring will also serve as a strong motivator for a new generation of yoga practitioners, while governments can leverage this data to design and implement promotive healthcare strategies and incentives to prevent diseases from reaching epidemic proportions [4]. Beyond individual practice, AI-based yoga solutions will stimulate data-driven research, fostering the development of disease-specific mobile applications that doctors can confidently recommend as part of their prescriptions. The ability to track yoga pose improvement and correlate it with disease lifecycle monitoring holds the promise of reducing the need for medication and surgical interventions. Furthermore, this AI tool can be instrumental in objectively evaluating yoga instructors and can serve as a valuable resource for addressing mental health issues within educational institutions, as well as in professional sectors such as the military and police forces. Ultimately, by popularizing yoga through accessible AI tools, we can significantly enhance society's resilience and ability to manage future pandemic situations.

**Use Case Status: Active, with ongoing deployment and iterative refinement.**

**Partners:** Ministry of AYUSH [6], yoga & wellness industry, medical practitioners, and academic institutes in the health domain.

## 2.2 Benefit of use case

AI can help bring affordable Healthcare solutions and ensure Good Health of all citizens – both on the curative and preventive side. Personalized healthcare solutions will become available to one and all. With availability of AI and Mobile based solution, Public health education in Yoga care can be implemented at population scale, reducing the load on hospitals. This AI tool can also be utilized for conducting examination of Yoga instructors in an objective manner, as well as a solution in mental health issues at school, college level as well as at organization level in sectors like military, police etc.

With popularization of Yoga, the ability to handle pandemic situations in the future will improve in the future.

## 2.3 Future Work

- Future work intends to expand the dataset to widen the scope of yoga poses and include detailed information about each. A mobile app can also be developed for ease of use by yoga practitioners.
- The development of standards is paramount. 'Landmarks' essential for the determination of yoga poses ought to be standardised from an interoperability perspective. These landmarks may subsequently be converted back into an Avatar-based yoga instructor. Initially, keypoints are extracted to delineate landmarks, which are then transmitted as input to the classifier model for pose prediction and accuracy estimation. Should the accuracy exceed seventy per cent, a timer shall commence counting the duration of the

pose; otherwise, a buzzer will sound, concurrently highlighting the landmarks in red. Furthermore, the vision models can be hosted locally to address any privacy-related concerns.

## 3 Use Case Requirements

- **REQ01**: It is critical to detect all 17 landmarks on the body, based on the input video stream of the Yoga posture.
- **REQ02:** It is critical to draw connections between them in white colour.
- **REQ03**: It is critical to preserve the privacy of user data by processing the same on the Mobile app itself.
- **REQ04**: It is critical that model should work with real-time videos of Yoga learners.
- **REQ05**: It is expected that if the current posture (real time) reaches an accuracy of threefold above 70% probability, then the colour turns green.
- **REQ06**: It is of added value to measure the time of holding the posture correctly.
- **REQ07**: It is of added value to notify the user by playing sound or music.
- **REQ08**: It is of added value that the mobile app is usable and mobile in perspective.
- **REQ09**: It is expected that the Mobile app provides a feature for collecting user feedback.
- **REQ10**: It is expected that the tracking and monitoring will also be a useful feature from an individual improvement perspective, as well as a research dimension.
- **REQ11**: It is expected that there should not be any lag, delay, or flickering in providing the feedback.
- **REQ12**: It is of added advantage that the Web-based component will help provide model updates and enhancements.
- **REQ13**: It is expected that a Security audit should be performed for the mobile & web application.

## 4 Sequence Diagram

- Sequence diagram

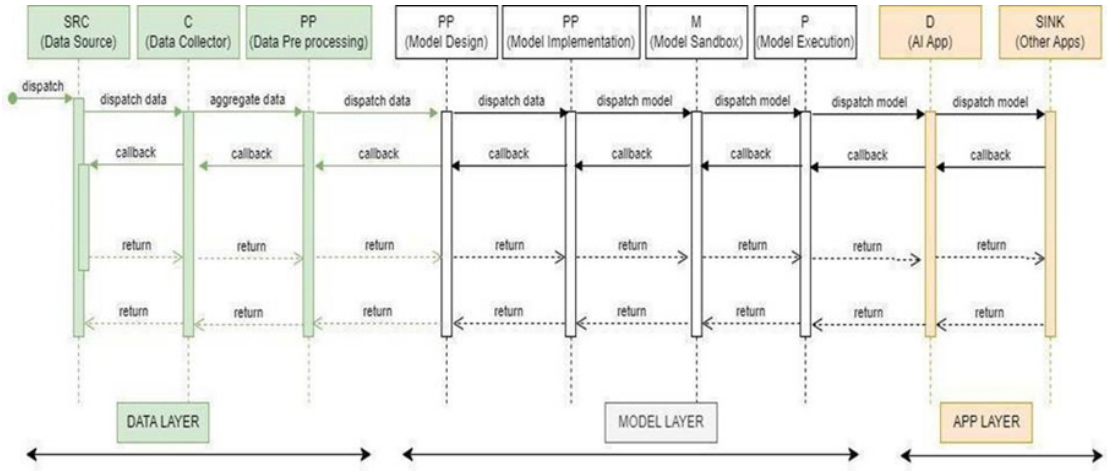



Figure 2.1: Sequence Diagram of our system

## 5 References

[1] M. Kumar, Yoga Poses Dataset GitHub Repository. Available: https://github.com/Manoj-2702/Yoga_Poses-Dataset.

[2] Yoga-82 Dataset. Available: https://sites.google.com/view/yoga-82/home.

[3] "Infinity Yoga tutor: Yoga posture detection and correction system," in *Proc. 2020 IEEE International Conference on Computing, Power and Communication Technologies (GUCON)*, 2020, Art. no. 9310832. Available: https://ieeexplore.ieee.org/document/9310832.

[4] Google, MoveNet SinglePose Thunder Model. Available: https://tfhub.dev/google/movenet/singlepose/thunder/4.

[5] TensorFlow.js, Import Keras Models Tutorial. Available: https://www.tensorflow.org/js/tutorials/conversion/import_keras.

[6] Ministry of Ayush, Government of India. Available: https://ayush.gov.in/.

# Use Case 3: CAFEIN – Federated Learning Platform

**Organization**: CERN

**Country**: Switzerland

**Contact Person(s)**:

Diogo Reis Santos, diogo.reis.santos@cern.ch,

Luigi Serio, luigi.serio@cern.ch

Alessandro Raimondo, a.raimondo@cern.ch

## 1 Use Case Summary Table

| Item | Details |
|---|---|
| Category | Healthcare |
| Problem Addressed | In many industries, sensitive data is distributed across different institutions. Traditional centralized machine learning approaches require data sharing that may compromise privacy, regulatory compliance, and data sovereignty. There is a growing need for solutions that enable collaborative analytics and AI model training without moving raw data. |
| Key Aspects of Solution | CERN's Federated Learning Infrastructure (CAFEIN) is a comprehensive, fully hosted platform—not just a framework—that provides an integrated software framework, robust software infrastructure, and dedicated server-side hardware infrastructure.<br>The platform is hosted, operated, and managed by Conseil Européen pour la Recherche Nucléaire (CERN), ensuring that it benefits from CERN's recognized network security, as well as its status as a non-profit international organization with non-military application principles.<br>CAFEIN facilitates secure federated learning and analytics by enabling decentralized model training while ensuring that sensitive data remains local. Advanced privacy-preserving technologies, such as secure aggregation and differential privacy, further protect data integrity and confidentiality. |
| Technology Keywords | Federated Learning, Federated Analytics, Federated Inference, Privacy-preserving, Machine Learning, Artificial Intelligence |
| Data Availability | Local data is maintained and secured by each participating organization, ensuring data privacy and regulatory compliance. |
| Metadata (Type of Data) | The platform is data-type agnostic. It supports structured and unstructured data, including numerical, text, and image data from various domains (e.g., sensor data, operational logs, imaging). |

(continued)

| Item | Details |
|---|---|
| Model Training and Fine-Tuning | – Supports federated analytics for privacy-preserving data analysis across decentralized datasets.<br>– Utilizes federated averaging techniques and secure aggregation protocols for decentralized training.<br>– Incorporates privacy-preserving enhancements (e.g., differential privacy) and supports heterogeneous model fine-tuning across data silos.<br>– Additionally, federated inference can be developed, enabling predictions without transferring data from heterogeneous sources. |
| Testbeds or Pilot Deployments | The platform is hosted, operated, and managed by CERN, ensuring robust infrastructure, security, and neutrality as part of its status as a non-profit international organization focused on non-military applications. |
| Code repositories | Not Available (not open sourced) |

# 2 Use Case Description

## 2.1 Description

CAFEIN is a cutting-edge federated learning platform developed at CERN. It addresses the challenge of training robust AI models in scenarios where data is distributed and sensitive. By enabling collaborative model training without centralising data, the platform enhances privacy, regulatory compliance, and data sovereignty. CAFEIN integrates state-of-the-art federated learning protocols with privacy-enhancing technologies, thus providing a scalable, secure, and efficient solution for multi-institutional AI projects. CAFEIN is a comprehensive, fully hosted platform—not just a framework—that provides an integrated software framework, robust software infrastructure, and dedicated server-side hardware infrastructure. The platform is hosted, operated, and managed by CERN, ensuring that it benefits from CERN's internationally recognized network security, as well as its status as a non-profit international organization with non-military application principles.

### Use Case Status: Operationally deployed, actively utilized, and iteratively optimized

Partners: CERN, World Health Organization (WHO) [8], the World Food Programme (WFP) [9], Fondazione Policlinico Universitario Agostino Gemelli Irccs (Italy) [10], Fundació Eurecat (Spain) [11], Fundació Hospital Universitari Vall D'Hebron – Institut de Recerca (Spain) [12], Katholieke Universiteit Leuven (Belgium) [13], Universität Basel (Switzerland) [14], and Siemens Healthineers AG (Germany) [15], among others.

## 2.2 Benefits of use case

Minimum Viable Product

- Good Health and Well-being – CAFEIN supports healthcare applications by enabling secure, decentralized analytics on sensitive health data, thereby improving medical research and patient care outcomes while maintaining data privacy.

- Industry, Innovation, and Infrastructure – By fostering innovative, privacy-preserving AI methods that can be deployed across sectors.
- Reduced Inequalities – By providing a robust federated learning platform that allows institutions from different regions and resource levels to collaborate without centralizing their data, CAFEIN helps bridge the digital divide and ensures equitable access to advanced AI technologies.
- Partnerships for the Goals – By enabling collaborative efforts between research institutions, industry, and government bodies without compromising data security.

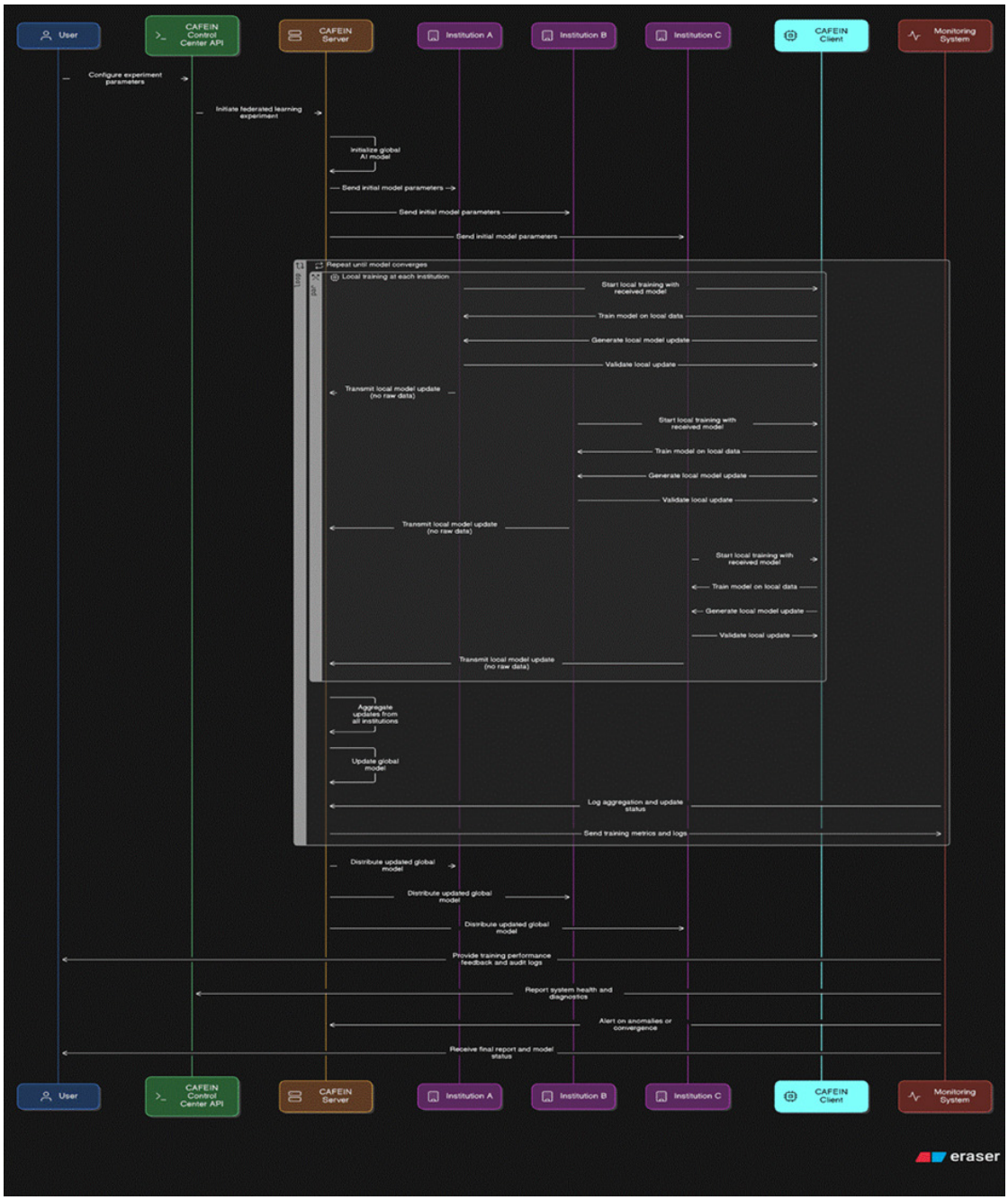
## 2.3 Future Work

- Ongoing EU Projects: Continue the deployment and integration of CAFEIN within ongoing EU projects to validate and refine its use in real-world scenarios.
- Expanded Collaborations: Strengthen and extend collaborations with key international organizations such as the World Health Organization (WHO) and the World Food Programme (WFP) to explore applications in public health and humanitarian initiatives.
- Federated Inference Deployment: Enhance the platform by deploying federated inference capabilities, enabling decentralized, real-time model predictions while ensuring data privacy remains uncompromised.
- Simplified Deployment: Focus on making the deployment process easier and more efficient for partner institutions, reducing technical barriers to adoption.
- Wider Deployment: Broaden the scope of deployment by targeting additional industry sectors and international research collaborations.
- Collaborative **Research:** Foster further collaborative research initiatives with academic and industry partners to innovate and optimize federated learning and inference techniques, ensuring CAFEIN remains at the forefront of privacy-preserving decentralized AI technology.

# 3 Use Case Requirements

- **REQ_01:** It is critical that the platform enable collaborative AI model training and analytics on distributed datasets without requiring the centralisation of raw data from participating institutions.
- **REQ_02:** It is critical that the platform ensure that sensitive data remains localized at each participating organization, preserving data privacy, supporting regulatory compliance (e.g., GDPR), and maintaining data sovereignty.
- **REQ_03:** It is critical that the platform is data-type agnostic, supporting structured and unstructured data, including numerical, text, and image data, to accommodate diverse research and industry domains.
- **REQ_04:** It is critical that the system provide capabilities for federated inference, allowing AI models to make predictions on decentralized data without transferring the raw input data to a central server.
- **REQ_05:** It is critical that the platform provides capabilities for federated inference, allowing AI models to make predictions on decentralised data without transferring the raw input data to a central server.
- **REQ_06:** It is critical that the platform implements comprehensive access controls, data encryption mechanisms (at-rest and in-transit for model updates), and tamper-evident audit logs to ensure system security and allow participants to verify against data leakage.

# 4 Sequence Diagram



## 5 References


[1] CAFEIN Project Website: https://cafein.web.cern.ch/

[2] A Carbon Tracking Model for Federated Learning: Impact of Quantization and Sparsification - https://arxiv.org/abs/2310.08087

[3] Brain MRI Screening Tool with Federated Learning - https://arxiv.org/pdf/2311.14086

[4] Feasibility Analysis of Federated Neural Networks for Explainable Detection of Atrial Fibrillation - https://doi.org/10.48550/arXiv.2410.19781

[6] Federated GNNs for EEG-Based Stroke Assessment - https://doi.org/10.48550/arXiv.2411.02286

[7] A Federated Learning Platform as a Service for Advancing Stroke Management in European Clinical Centers - https://arxiv.org/abs/2410.13869

[8] CERN, "CAFEIN – CERN AI Federated Infrastructure," [Online]. Available: https://cafein.web.cern.ch/.

[9] World Health Organization (WHO), "About WHO," [Online]. Available: https://www.who.int/about.

[10] World Food Programme (WFP), "About WFP," [Online]. Available: https://www.wfp.org/about.

[11] Fondazione Policlinico Universitario Agostino Gemelli IRCCS, "Istituto di Ricovero e Cura a Carattere Scientifico," [Online]. Available: https://www.policlinicogemelli.it/.

[12] Eurecat – Fundació Eurecat, "Centre Tecnològic de Catalunya," [Online]. Available: https://eurecat.org/.

[13] Vall d'Hebron Research Institute (VHIR), "Institut de Recerca," [Online]. Available: https://www.vhir.org/.

[14] Katholieke Universiteit Leuven, "KU Leuven – Inspiring the Outstanding," [Online]. Available: https://www.kuleuven.be/.

[15] Universität Basel, "University of Basel," [Online]. Available: https://www.unibas.ch/.

[16] Siemens Healthineers AG, "Official Website," [Online]. Available: https://www.siemens-healthineers.com/.

# Use Case 4: Smartphone-over-Microscope Diagnosis: Empowering Community Health Workers with TinyML for Malaria Parasite Detection in Rural Areas

**Organization**: CSIR - Institute for Scientific and Technological Information

**Country**: Ghana

Bolgatanga Technical University

**Contact Person(s)**:

Dennis Agyemanh Nana Gookyi dennisgookyi@gmail.com
Fortunatus Aabangbio Wulnye fortunatuswulnye@outlook.com
Michael Wilson yboabengwilson@gmail.com
Roger Kwao Ahiadormey rogerkwao@gmail.com
Moses Abambire abambiremoses@gmail.com
Paul Danquah pauldanquah@yahoo.com
Raymond Gyaang gyaangraymond@outlook.com

## 1 Use Case Summary Table

| Item | Details |
|---|---|
| Category | Healthcare |
| Problem Addressed | The primary issue addressed is the lack of reliable and timely malaria diagnosis in rural communities, where access to skilled laboratory technicians and diagnostic facilities is limited. Traditional microscopy requires trained professionals and laboratory infrastructure, while Rapid Diagnostic Tests (RDTs) can be less accurate and prone to false negatives. This gap in diagnostic capability hampers effective malaria treatment and control efforts in underserved areas. |

(continued)

| Item | Details |
|---|---|
| Key Aspects of Solution | • Smartphone-Mounted Microscopy: Use of a smartphone positioned over a microscope eyepiece to capture high-resolution images of blood slides for analysis<br>• TinyML-Powered Real-Time Image Analysis: Deployment of lightweight AI models on the smartphone to detect malaria parasites (Trochozoites) and White Blood Cells (WBCs) directly from captured images<br>• Training with Labeled Dataset: Use of the Lacuna Malaria Detection Challenge dataset to train and calibrate the AI model for accurate classification and detection<br>• Portable and Accessible Diagnostics: Providing community health workers with an easy-to-use, portable tool that enhances malaria diagnosis without the need for full laboratory setups |
| Technology Keywords | TinyML, Smartphone Microscopy, Malaria Detection, Deep Learning, Edge AI, Trochozoite Identification, WBC Detection |
| Data Availability | Public [1] |
| Metadata (Type of Data) | Visual data (high-resolution images of blood slides) |
| Model Training and Fine-Tuning | The AI model is trained using supervised learning techniques, specifically deep learning with convolutional neural networks (CNNs), to classify and detect malaria parasites (Trochozoites) and WBCs from blood slide images. Fine-tuning is performed using labeled visual data from the Lacuna Malaria Detection Challenge dataset to improve accuracy and adapt the model for real-world variability in image quality. TinyML frameworks are then used to optimize the model for efficient deployment on resource-constrained smartphone devices. |
| Testbeds or Pilot Deployments | Lacuna Malaria Detection Challenge [2] |
| Code repositories | Smartphone-over-Microscope-Diagnosis [3] |

# 2 Use Case Description

## 2.1 Description

This use case centers on utilizing a smartphone mounted over a microscope to improve the diagnosis of malaria parasites in rural communities. The approach taps into TinyML's capabilities for real-time image analysis, thereby empowering community health workers with a portable and efficient diagnostic tool. By leveraging a smartphone to capture high-resolution images of blood slides through the microscope eyepiece, the system aims to facilitate accurate and timely detection of malaria infections, specifically identifying Trochozoites and WBCs.

The problem this use case addresses is limited access to reliable and timely malaria diagnosis in remote areas, where resources and trained professionals may be scarce. Current solutions, such as traditional microscopy or RDTs, often face limitations. Conventional microscopy requires

skilled technicians and may not be feasible in rural settings, while RDTs can yield false negatives and are less accurate.

The dataset employed for training the model is sourced from the Lacuna Malaria Detection Challenge, which contains diverse and labeled images of blood slides, essential for calibrating the algorithm for effective detection. Utilizing this dataset ensures a robust model capable of distinguishing between the relevant classes—Trochozoites and WBCs.

The benefits of this proposed solution include enhanced accessibility to malaria diagnosis, reduced dependency on laboratory facilities, and the utilization of familiar technology (smartphones) by healthcare workers. However, potential drawbacks include the need for adequate smartphone hardware and battery life, which could affect performance in areas with unreliable electricity supply. Additionally, while the model can achieve high accuracy with sufficient training data, variability in image quality and environmental factors may still pose challenges.

In conclusion, the smartphone-over-microscope approach combines modern technology with community health efforts to improve malaria detection significantly, ultimately aiming for better health outcomes in underserved populations.

## 2.2 Benefits of the use case

By leveraging AI technology, this approach enhances diagnostic accuracy and accessibility, making it a vital tool in combating malaria, a leading cause of morbidity and mortality in many developing regions.

AI-driven image analysis automates the detection of malaria parasites, ensuring timely diagnoses that can significantly improve patient outcomes and reduce the burden on healthcare systems. This helps to expand access to essential health services and promoting universal health coverage, particularly in underserved areas. Empowering community health workers with AI tools eliminates barriers caused by shortages of trained medical professionals and offers a scalable solution to monitoring and managing malaria outbreaks.

Furthermore, the technological innovation represented by this use case fosters resilient infrastructure and promots research and technological development. As AI models are trained on datasets like the Lacuna Malaria Detection Challenge, they can evolve to support various health initiatives beyond malaria, thus enhancing local health capacities.

In summary, the integration of AI in malaria detection not only addresses immediate health challenges but also contributes to broader efforts in innovation, infrastructure development, and improving health equity.

### Partners

Bolgatanga Technical University: Responsible for collecting additional blood slide samples from local healthcare facilities to expand and diversify the dataset. Their role ensures the AI model is trained on region-specific data, improving its accuracy and robustness for real-world deployment in rural communities.

## 2.3 Future Work

Future work for this project includes expanding the blood slide dataset by collaborating with more local clinics through Bolgatanga Technical University, ensuring the model has a diverse and representative set of images for better generalization. The AI model will be further improved by incorporating image augmentation techniques to address issues related to varying lighting and image quality conditions. Additionally, a user-friendly mobile application interface will be developed to make the smartphone-microscope setup easier to operate by community health workers with minimal training.

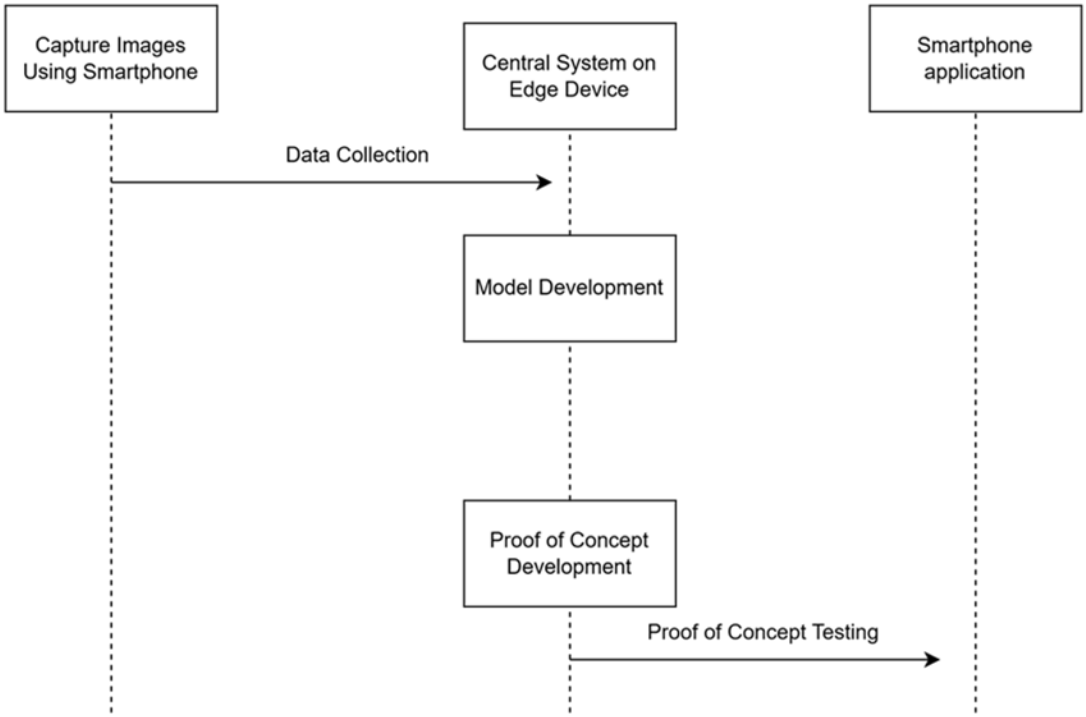
To support these enhancements, upgraded smartphones with better camera resolution and processing power will be required to handle more complex TinyML models. More microscopes with adaptable mounts for stable smartphone imaging will also be needed, along with funding for ongoing field data collection, device maintenance, and training programs. Cloud storage solutions will be essential for securely backing up and managing large datasets.

In terms of collaborations and expansions, there are plans to work with regional healthcare organizations and NGOs to pilot the system in more rural communities, potentially broadening its reach. The AI system could also be expanded to detect other blood-borne diseases, such as sickle cell anemia or bacterial infections, further improving healthcare in remote areas. Additionally, partnerships with technology companies or research institutions will be explored to optimize the TinyML models and investigate more efficient deployment on low-cost devices.

# 3 Use Case Requirements

- **REQ-01:** Smartphone with high-resolution camera capability and sufficient processing power to run TinyML models.
- **REQ-02:** Optical microscope compatible with smartphone mounting for clear blood slide imaging.
- **REQ-03:** Trained TinyML model optimized for real-time malaria parasite (Trochozoite) and WBC detection.
- **REQ-04:** Reliable power sources (e.g., battery banks or solar chargers) to support smartphone use in areas with unstable electricity.
- **REQ-05:** Secure data storage and management system for saving and labeling new blood slide images collected during fieldwork.
- **REQ-06:** Basic training materials and sessions for community health workers on operating the smartphone-microscope setup and interpreting results.
- **REQ-07:** Collaboration framework with Bolgatanga Technical University for continuous dataset expansion and validation.

## 4 Sequence Diagram




## 5 References

[1] Lab, Makerere AI, "Lacuna Malaria Datasets." Harvard Dataverse, 2023. doi: 10.7910/DVN/VEADSE. Available: https://dataverse.harvard.edu/citation?persistentId=doi:10.7910/DVN/VEADSE. [1] "CellScope – UC Berkeley." Available: https://cellscope.berkeley.edu/.

[2] Zindi, "Lacuna Malaria Detection Challenge," 2024. Available: https://zindi.africa/competitions/lacuna-malaria-detection-challenge.

[3] D. Gookyi, "Smartphone-over-Microscope-Diagnosis," GitHub, May 29, 2025. Available: https://github.com/dennisgookyi/Smartphone-over-Microscope-Diagnosis. [Accessed: Jun. 09, 2025]

[4] World Health Organization, "Malaria microscopy quality assurance manual," 2016. Available: https://www.who.int/publications/i/item/9789241549394.

[5] Rojas, A., & Rojas, A. (2017). Malaria Diagnosis: A Review of the Current State and Future Perspectives. Journal of Parasitology Research, 2017, 1-10. Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5308250/

[6] Raju, G., Ranjan, A., Banik, S. et al. A commentary on the development and use of smartphone imaging devices. Biophys Rev 16, 151–163 (2024). https://doi.org/10.1007/s12551-023-01175-1.

# Use Case 5: AI-Driven Biometric Animal Registration for Animal Welfare and Global Health

**Organization:** iSciLab Corporation

**Country**: Republic of Korea

**Contact person(s)**:

Changhyun Byun, Ph.D, cbyun@iscilab.com

Kiyoung Jang, ky.jang@kra.co.kr

## 1 Use case summary table

| Item | Details |
|---|---|
| Category | Healthcare |
| The Problem to be addressed | Traditional animal registration using Radio Frequency Identification (RFID) microchips is invasive, costly, and impractical in many regions. This limits pet identification, hinders rabies vaccination tracking, and results in poor enforcement of animal welfare policies. Especially in low-income areas, unregistered stray animals contribute to zoonotic disease spread and animal neglect. |
| Key Aspects of the Solution | • AI-powered biometric animal identification (e.g., nose print, facial recognition) using mobile devices<br>• Non-invasive, low-cost alternative to RFID chips<br>• Digital vaccination and health record management<br>• Expansion to livestock (e.g., cattle, horses)<br>• Supporting global deployment and international standardization in collaboration with ITU-T to ensure scalable and interoperable AI solutions. |
| Technology keywords | Biometric Recognition, AI for Animal Welfare, Digital Identity, Vaccination Tracking |
| Data availability | • Dog nose-print image data is available online [1]<br>• Cattle nose-print image data is available online [2]<br>• Horse biometrics image data is available online [3] |
| Metadata (type of data) | Animal biometric images (nose print, facial features), species/breed/age/owner info, vaccination history |
| Model Training and Fine-Tuning | Models trained on labeled biometric image datasets; transfer learning and fine-tuning for species-specific accuracy |
| Testbeds of pilot deployments | • Pilot in Korea in collaboration with Korea Racing Authority (KRA) for digital horse passport<br>• Field testing in African regions for rabies vaccination tracking and pet ID<br>• NGO and government partnerships for wider rollout |

(continued)

| Item | Details |
|---|---|
| Code repositories | The source code and related datasets supporting this use case are publicly available and can be accessed through references |

# 2 Use case description

## 2.1 Description

Traditional animal registration using RFID microchips is invasive, costly, and prone to issues such as chip migration, incompatibility, or failure. These barriers hinder widespread adoption, especially in low-resource settings, resulting in unregistered animals and poor vaccination tracking—an acute public health risk, particularly for stray animals in regions like Africa. Additionally, physical identifiers like tags or microchips are vulnerable to tampering or misuse, leading to animal identity fraud, where one animal is intentionally or mistakenly presented as another. This undermines the integrity of registration, traceability, and disease control systems.

To overcome these challenges, a non-invasive AI-powered biometric system has been developed using canine nose prints. These nose patterns are unique and stable from as early as two months of age [4], [5]. Deep learning models, trained in collaboration with veterinary experts, enable instant, contactless identification via smartphones—eliminating the need for surgical implants or specialized scanners.

This solution offers a scalable, cost-effective alternative to microchips. It enhances accuracy, prevents duplicate or fraudulent registrations—including identity fraud—and simplifies data updates through mobile apps. Crucially, it supports better disease tracking and aligns with global health initiatives such as WHO's "Zero by 30" [6] rabies elimination goal—making animal management systems safer, more humane, and accessible worldwide.

### Use case status: The use case is part of a larger product development

This use case is part of a global effort to develop an AI-based biometric animal ID system, with pilots in South Korea, Australia, and Ghana showing its ease of use and highlighting the need for broader integration and multi-species support.

**Partner**: Korea Racing Authority (KRA) [7]

## 2.2. Benefit of use case

- Good Health and Well-being Systematic registration and tracking of animals enhance disease surveillance and prevention, reducing the risk of zoonotic diseases (e.g., rabies). This contributes to better public health and improves the well-being of both animals and humans.
- Implementing biometric animal registration fosters technological advancements in pet tech, AI-driven pet healthcare, and smart animal management systems, promoting innovation in the pet industry.
- By encouraging responsible pet ownership, biometric registration helps combat illegal breeding and unethical animal trade.

- By promoting responsible pet ownership, it contributes to the welfare of domestic animals and supports broader wildlife conservation efforts, including anti-poaching initiatives.
- A standardized biometric animal identification system strengthens policies against animal abuse, illegal pet trade, and pet abandonment. It enables effective law enforcement and enhances animal welfare regulations.

### 2.3 Future work

With support from this program, we aim to develop and deploy an AI-powered biometric animal registration system that uses technologies such as nose print and facial recognition to enable non-invasive, smartphone-based animal identification. This approach provides a more accessible and cost-effective alternative to traditional identification methods like RFID chips, ear tattoos, and ear-notching, which are often invasive, expensive, and difficult to implement—particularly in low-income regions where such approaches can be logistically and financially prohibitive. The system will initially focus on companion animals, with features that include tracking vital health information like rabies vaccinations—supporting initiatives such as the WHO's "Zero by 30" campaign to eliminate dog-mediated human rabies deaths by 2030.

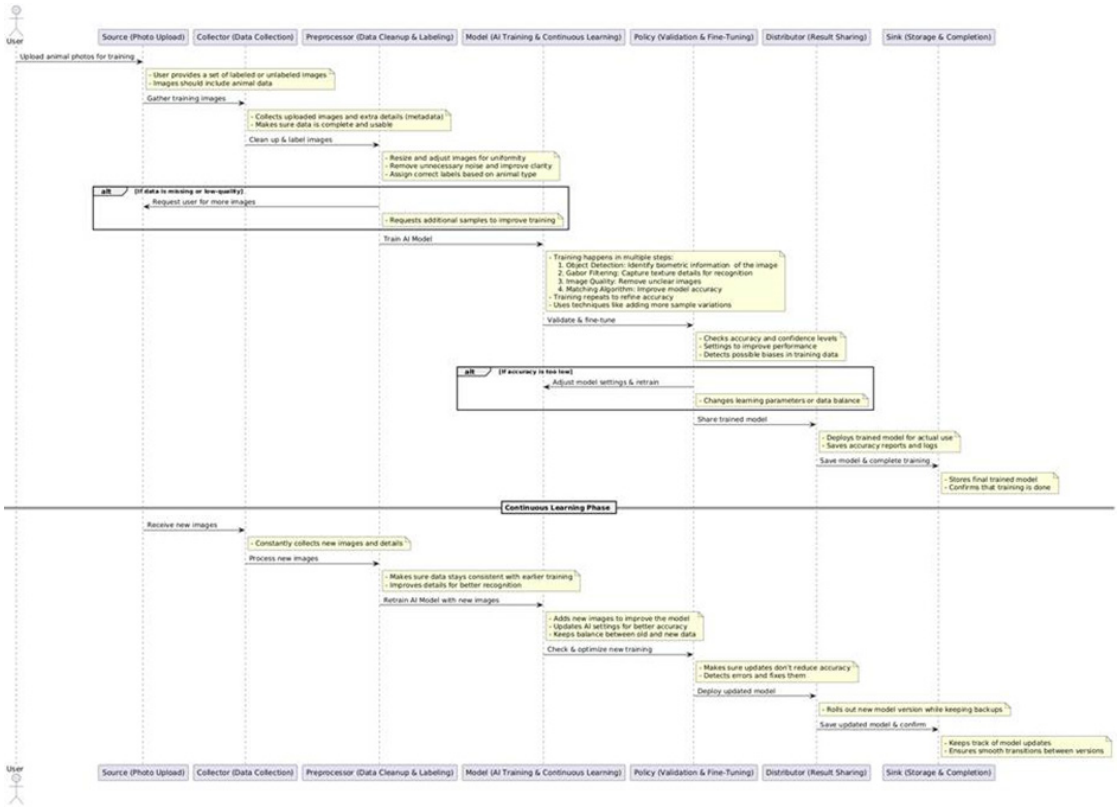
Our immediate priority is to launch a robust mobile platform and pilot it in regions where stray and unregistered animals pose public health risks. The system will later be expanded to livestock such as cattle and horses to support broader needs in animal traceability, food safety, and disease management. In collaboration with the Korea Racing Authority (KRA), we will also develop a digital equine passport system that integrates biometric identification with health and performance data to benefit the racing and breeding industries.

To drive global adoption, we plan to work with governments, NGOs, and international standardization bodies. A key milestone will be the proposal of a new work item in ITU-T Study Group 17 to expand the existing x.pet_auth framework [8] to cover livestock, enabling standardized and interoperable animal identity systems worldwide. Through this initiative, we aim to improve global animal welfare, strengthen public health systems, and promote inclusive digital transformation in veterinary services.

## 3 Use case requirements

- **REQ-01:** The solution must support AI-based biometric recognition (nose pattern, face landmarks, fur pattern) for various species including dogs, cattle, and horses.
- **REQ-02:** The solution must be usable via standard mobile devices (e.g., smartphones), without the need for specialized scanners.
- **REQ-03:** The system must enable digital storage and retrieval of comprehensive animal metadata, including identification details, vaccination history, and health information.
- **REQ-04:** The system must allow secure owner-pet identity linkage, ensuring accurate identification and updates.
- **REQ-05:** The system should comply with global standards for animal identification and data interoperability.
- **REQ-06:** The solution must support offline registration and sync when connectivity is restored, especially for rural deployments.

## 4 Sequence diagram




## 5 References

[1] ISCI Lab, "Noseprint Public Data (Dogs)," GitHub repository, 2022. [Online]. Available: https://github.com/iscilab-repo/noseprint_public_data. [Accessed: 6-Jun-2025].

[2] ISCI Lab, "Cow Noseprint Dataset," GitHub repository, 2022. [Online]. Available: https://github.com/iscilab-repo/cow-noseprint-data. [Accessed: 6-Jun-2025].

[3] AI Hub Korea, "AI Noseprint Image Dataset for Animal Identification," AI Hub, 2024. [Online]. Available: https://www.aihub.or.kr/aihubdata/data/view.do?dataSetSn=71707. [Accessed: 6-Jun-2025].

[3] Choi, H.I.; Kim, M.-Y.; Yoon, H.-Y.; Lee, S.; Choi, S.S.; Han, C.Y.; Moon, H.P.; Byun, C.; Kwon, S.-H. Study on the Viability of Canine Nose Pattern as a Unique Biometric Marker. Animals 2021, 11, 3372. https://doi.org/10.3390/ani11123372

[4] Choi, H.I.; Lee, Y.; Shin, H.; Lee, S.; Choi, S.S.; Han, C.Y.; Kwon, S.-H. The Formation and Invariance of Canine Nose Pattern of Beagle Dogs from Early Puppy to Young Adult Periods. Animals 2021, 11, 2664. https://doi.org/10.3390/ani11092664

[5] WHO Zero by 30 Strategy. Available online: https://www.who.int/initiatives/zero-by-30

[6] Korea Racing Authority – https://www.kra.co.kr/globalEn/

[7] [ITU-T X.1095] Recommendation ITU-T X.1095(2022) Entity authentication service for pet animals using telebiometrics. Available online: https://www.itu.int/itu-t/workprog/wp_item.aspx?isn=18003

# Use Case 6: AI-Powered Dental CAD Solution for Fully Automated Digital Prosthesis Design

**Organization**: Imagoworks Inc.

**Country**: Republic of Korea

**Contact Person(s)**:

Hoseong Shim, hsshim@imagoworks.ai

Junsik Yoon, 3DmeJS@imagoworks.ai

## 1 Use Case Summary Table

| Item | Details |
|---|---|
| Category | Healthcare |
| Problem Addressed | Dental professionals and technicians continue to face significant challenges in prosthesis design, as traditional dental CAD workflows remain labor-intensive, requiring a high level of manual intervention and expertise. These processes are not only time-consuming but also prone to inconsistencies and human error, which can negatively impact treatment quality, extend patient discomfort, and increase clinical risks due to poorly fitted or inaccurately designed prostheses. This highlights the pressing need for an automated, standardized solution that can enhance accuracy, speed, and consistency in dental restoration. |
| Key Aspects of Solution | • Automated prosthesis design using AI to reduce manual work and speed up treatment.<br>• AI-driven batch processing system to enable high-volume prosthesis production<br>• Cloud-based AIaaS model for accessible, scalable, and cost-efficient Computer-Aided Design (CAD) solutions |
| Technology Keywords | • AIaaS<br>• Deep Learning,<br>• Cloud Computing<br>• Web based 3D CAD<br>• 3D Data Processing<br>• AI-based Design Automation<br>• Digital Dentistry Solutions |
| Data Availability | The dataset supporting this use case is publicly available and can be accessed at [1]. |
| Metadata (Type of Data) | Dental 3D scanned data (Images) |

(continued)

| Item | Details |
|---|---|
| Model Training and Fine-Tuning | • Data preparation: Standard Tessellation Language (STL) preprocessing, noise filtering, anatomical labelling.<br>• Model architecture: Convolutional Neural Networks (CNNs) and Generative Adversarial Networks (GANs) for 3D mesh analysis. (PyTorch, TensorFlow)<br>• Training: GPU-based batch training with data augmentation and dental-specific loss functions.<br>• Hyperparameter tuning: Optimization via Optuna; cross-validation with clinical datasets.<br>• Fine-tuning: Continuous learning from real-world usage; transfer learning for specific cases.<br>• Deployment: Cloud-based AIaaS; model tracking and update via MLflow. |
| Testbeds or Pilot Deployments | The system has been evaluated in real-world settings through pilot deployments, demonstrating its practical feasibility and user acceptance [2], [3], [4], [5]. |
| Code repositories | Private |

# 2 Use Case Description

## 2.1 Description

Dental professionals and technicians continue to face significant challenges in prosthesis design, even as digital technologies become more prevalent in clinical and laboratory environments. Traditional dental CAD workflows remain labour-intensive, requiring a high level of manual intervention and expertise.

These workflows are not only time-consuming but also prone to inconsistencies and human error, which can negatively impact treatment quality, extend patient discomfort, and increase clinical risks due to poorly fitted or inaccurately designed prostheses. This demonstrates the pressing need for an automated, standardized solution that can enhance accuracy, speed, and consistency in dental restoration.

Dentbird Solutions introduces an AI-powered dental CAD system designed to address these limitations by automating the most time-consuming steps in prosthesis design. Leveraging advanced deep learning and 3D data processing technologies, the software automatically classifies upper and lower jaw data, detects margin lines, and generates patient-specific prosthesis (crown, implant crown, bridge, inlay, and onlay) designs in a matter of minutes.

The system also supports continuous mass production workflows, significantly alleviating the workload of dental technicians. As a cloud-based AI-as-a-Service (AIaaS) software, Dentbird eliminates the need for local software installation, making advanced dental CAD technology accessible across various clinical and laboratory settings.

Over time, the AI continuously improves through user feedback and new clinical data, ensuring optimal performance across a wide range of restorative cases. By streamlining workflows and minimizing variability, Dentbird Solutions enables faster, more predictable treatment outcomes—ultimately contributing to the advancement of global oral healthcare standards.

**Partners:**

- Ci Medical Co.,Ltd. (Japan)
- Medit Corp. (Korea)
- Yonsei University College of Dentistry (Korea)
- Gyeongnam Dental Care Center for Persons with Special Needs (Korea)
- Jeonbuk National University School of Dentistry (Korea)
- Kyunghee University College of Dentistry (Korea)
- CIMsystem s.r.l. (Italy)
- ASIGA (Australia)
- ModuleWorks GmbH (Germany)

## 2.2 Benefits of the use case

This use case significantly improves access to high-quality oral healthcare. Through the application of AI-driven diagnostic tools, automated prosthetic design, and standardized digital workflows, the solution reduces treatment time, enhances accuracy, and lowers costs for both patients and providers. This ensures that a greater number of individuals can receive timely and precise dental care, ultimately improving global oral health outcomes. Furthermore, by automating labor-intensive tasks, the solution enhances operational efficiency for dental professionals, supporting improved work-life balance, increased profitability, and better overall well-being within the dental workforce.

This use case introduces groundbreaking innovation in the field of digital dentistry. Dentbird Solutions is the first in the industry to integrate Generative AI with 3D imaging and dental CAD technology, revolutionizing traditional workflows. It further advances industry infrastructure by adopting a cloud-based AI-as-a-Service (AIaaS) model, which lowers the barrier to entry for digital dental tools and expands accessibility across clinics and labs. This new model not only supports scalable growth of dental practices but also stimulates industrial advancement and the creation of new job opportunities in AI, software, and digital healthcare sectors.

## 2.3 Future Work

To enhance 3D design for dental prosthesis based on Generative AI, it is essential to gather high-quality training data. This will serve as the foundation for training and improving AI models, and securing new high-quality data is crucial for expanding AI models and functionalities.

The allocated resources will be utilized to collect diverse datasets, particularly from complex or rare cases, to improve AI accuracy.

Based on the collected data, AI training will be conducted, and continuous innovation efforts will be made to meet growing market demands and user expectations. Existing AI models will be improved, and new models will be developed to ensure the solution remains effective and stable across a wider range of treatment cases.

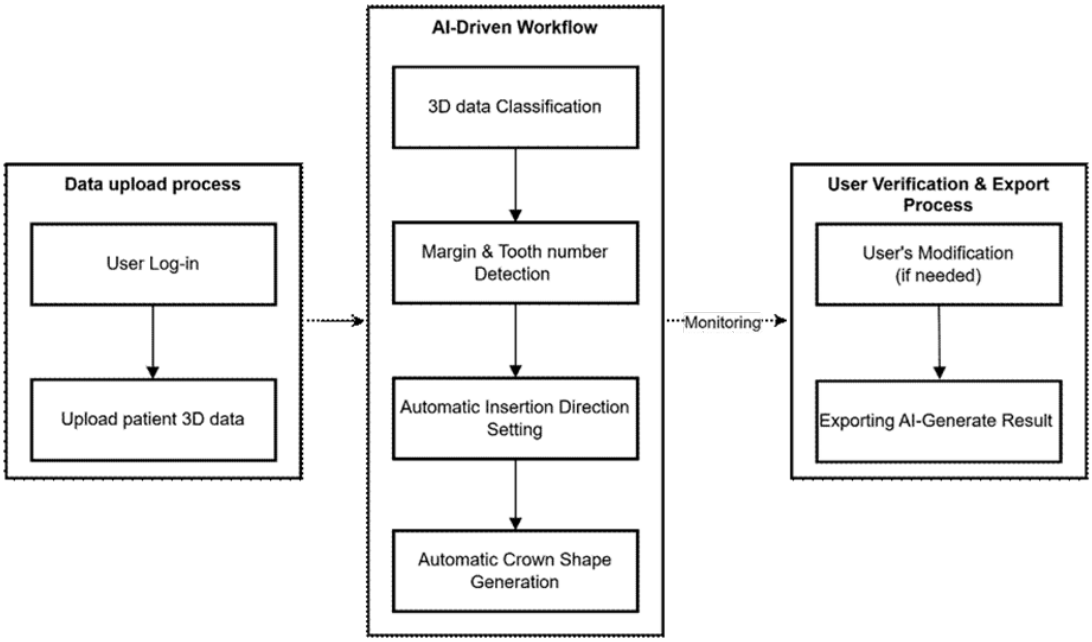
Creation of New Variations and Expansions for Clinical Applications: New AI models will be developed to expand their clinical applications in dentistry. This includes launching new solutions for dental prosthesis that have not yet been addressed.

In addition to supporting single crowns, bridges, and inlays, new solutions will be developed and launched for Full-mouth restorations, Implant surgical guides, and Automated design of custom implant abutments. Through these advancements, market share will be increased, and contributions will be made to improving global oral health.

## 3 Use Case Requirements

- **REQ-01:** It is critical that the Dentbird Solution collects high-quality data, including complex and rare clinical cases, to ensure sufficient diversity and precision in AI training and design performance.
- **REQ-02:** It is essential that the platform supports ongoing model development and optimization using deep learning techniques, enabling accurate crown and prosthesis designs across a wide range of patient scenarios.
- **REQ-03:** It is required that the solution expands clinical applicability by supporting implant surgical guides, and custom abutment designs, thereby increasing the utility of AI in diverse prosthetic treatments.
- **REQ-04:** It is critical that the system incorporates user feedback into continuous model updates, refining AI performance and enhancing user experience through data-driven improvement cycles.

## 4 Sequence Diagram




## 5 References

[1] Imagoworks Inc., "Dentbird: AI-powered dental CAD solution," Dentbird, [Online]. Available: https://www.dentbird.com/. [Accessed: 6-Jun-2025].

[2] J.-H. Cho, G. Çakmak, Y. Yi, H.-I. Yoon, B. Yilmaz, and M. Schimmel, "Tooth morphology, internal fit, occlusion and proximal contacts of dental crowns designed by deep learning-based dental software: A comparative study," *J. Dent.*, vol. 141, 104830, 2024. [Online]. Available: https://doi.org/10.1016/j.jdent.2023.104830

[3] J.-H. Cho, G. Çakmak, J. Choi, D. Lee, H.-I. Yoon, B. Yilmaz, and M. Schimmel, "Deep learning-designed implant-supported posterior crowns: Assessing time efficiency, tooth morphology, emergence profile, occlusion, and proximal contacts," *J. Dent.*, vol. 147, 105142, 2024. [Online]. Available: https://doi.org/10.1016/j.jdent.2024.105142

[4] G. Çakmak, J.-H. Cho, J. Choi, H.-I. Yoon, B. Yilmaz, and M. Schimmel, "Can deep learning-designed anterior tooth-borne crown fulfill morphologic, aesthetic, and functional criteria in clinical practice?" *J. Dent.*, vol. 150, 105368, 2024. [Online]. Available: https://doi.org/10.1016/j.jdent.2024.105368

[5] J. Choi, J. Ahn, and J.-M. Park, "Deep learning-based automated detection of the dental crown finish line: An accuracy study," *J. Prosthet. Dent.*, vol. 132, no. 6, pp. 1286.e1–1286.e9, Dec. 2024. [Online]. Available: https://doi.org/10.1016/j.prosdent.2023.11.018

# Use Case 7: AI-Powered Smart Healthcare for Remote & Underserved Communities

**Organization**: Tata Elxsi

**Country**: India

**Contact Person(s)**:

Jesu Maria Antony, jesumariaantony.s@tataelxsi.co.in

Bhuvaneshwari S, bhuvaneshwari.s@tataelxsi.co.in

Rajesh M, rajesh.m2@tataelxsi.co.in

## 1 Use Case Summary Table

| Item | Details |
|---|---|
| Category | Healthcare |
| Problem Addressed | Many rural and underserved communities face the critical challenge of limited access to quality healthcare. This isn't just an inconvenience; it's a systemic issue driven by a combination of inadequate infrastructure, significant geographic isolation, and persistent healthcare workforce shortages. As a result, residents in these areas often experience delayed diagnoses, limited treatment options, and poorer health outcomes. |
| Key Aspects of Solution | • Real-time data exchange<br>• AI-enabled diagnostics and decision support<br>• Remote monitoring and mobile health services<br>• Personalized, context-aware healthcare delivery |
| Technology Keywords | Remote Healthcare, Intelligent Health Systems, AI Diagnostics, Continuous Monitoring, Connectivity, Preventive Health |
| Data Availability | Combination of public health data and private, anonymised patient data from pilot regions |
| Metadata (Type of Data) | Heartbeat in Beats per minute (bpm), sleep duration in hours (hr), skin temperature in degrees Celsius |
| Model Training and Fine-Tuning | Isolation Forest models for anomaly detection and Random Survival Forests for survival analysis are trained on anonymised health datasets to detect early signs of illness, recommend treatments, and personalise health plans |

(continued)

| Item | Details |
|---|---|
| Testbeds or Pilot Deployments | Minimum Viable Product (MVP) in initial discussions/development with features around<br>1 Real-Time Health Data Collection (Controlled Environment)<br>2 Remote Patient Monitoring & Shadow Mode Streaming<br>3 AI Diagnostics & Decision Support<br>Autonomous Triage System (Shadow Mode) |
| Code repositories | Not Available |

# 2 Use Case Description

## 2.1 Description

This initiative introduces an AI-driven healthcare model to tackle the persistent healthcare disparities in remote and underserved areas. These disparities typically arise from inferior infrastructure, a shortage of clinicians, and inadequate connectivity. Traditional healthcare systems often falter in such environments, leading to delayed treatment and an increase in mortalities. Coupled with ultra-reliable, low-latency communications, this new solution aims to bridge the critical care gap. By leveraging mobile health units and wearable/remote sensors, the framework is designed to enable real-time diagnosis, real-time monitoring, and remote consultations. This approach effectively transforms healthcare delivery in resource-limited environments. The framework's core aims are clear: to significantly improve healthcare access and the quality of care in underserved communities, to encourage early detection and intervention for both chronic and acute conditions, and to alleviate the burden on central hospitals and emergency services. This innovative solution integrates an AI-supported health advisor for diagnosis and decision support, utilises intelligent kits and wearables for remote patient monitoring, and employs context-aware personalisation based on environmental and patient data. Furthermore, it establishes real-time communication channels between patients, local healthcare workers, and urban specialists, ensuring timely and expert medical advice. The use of digital twin technology, creating dynamic virtual patient representations from live data, coupled with AI models, allows for continuous monitoring, predictive risk assessment, and enhanced clinical decision-making. Deployment will follow a phased strategy, beginning with onsite validation in controlled clinic environments to test sensor integration and AI model performance, progressing to hybrid deployment with remote monitoring supported by nearby clinics, and culminating in full remote deployment with edge AI for autonomous triage. The selection of wearables is modular and tailored to specific conditions, including general health monitoring, maternal and infant care, chronic disease management, and neurological or elderly care. Adaptive AI model training will incorporate techniques like transfer learning, federated learning, and modular design to ensure efficiency and privacy, while model explainability mechanisms such as feature attribution, visual heatmaps, and confidence scores will be crucial for building clinician trust and ensuring transparent decision-making. The user interface will be multimodal and context-aware, featuring chat-based AI assistants, mobile dashboards for clinicians, and support for various data inputs, including text, voice, image, video, and sensor streams. The anticipated impact of this AI-powered framework is substantial, promising

expanded healthcare reach to marginalised populations, improved patient outcomes through early detection and remote interventions, empowerment of local caregivers with AI support, and a reduction in overall healthcare delivery costs over time. This comprehensive approach seeks to provide scalable, context-aware healthcare with minimal reliance on physical infrastructure, fundamentally changing how healthcare is accessed and delivered in underserved communities

**Use Case Status:** MVP in initial discussions/development

**Partners** Tata Elxsi - Lead implementer and system designer [5]

## 2.2 Benefits of use case

The use case enables the following benefits:

- Improve healthcare access and quality in underserved areas
- Enable early detection and intervention for chronic and acute conditions
- Reduce pressure on centralized hospitals and emergency services
- Provide scalable, context-aware healthcare with minimal physical infrastructure

## 2.3 Future Work

Future works will extend to

- Expanding coverage to more rural and underserved regions
- Incorporate mental health and maternal care modules
- Develop multilingual voice assistants for accessibility
- Interactive Explainability Dashboards: For clinicians to explore "why" behind alerts and tweak sensitivity
- Federated Feedback Loops: Feedback from clinicians in real-world settings can fine-tune and validate explanation quality.
- Integrate with national health databases and emergency alert systems
- Explore a partnership with telecom operators for optimised service delivery
- Pursue validation studies and regulatory pathways for clinical acceptance

# 3 Use Case Requirements

- **REQ-01:** It is critical that a real-time Health Data Collection in a Controlled Environment is implemented.
- **REQ-02:** It is critical that a Health Data Streaming in Field (Shadow Mode).
- **REQ-03:** It is critical that the system is an autonomous health data acquisition and escalation system.
- **REQ-04:** It is critical that AI diagnostic model testing be done in a controlled environment.
- **REQ-05:** It is critical that AI triage Inference is done in shadow mode.
- **REQ-06:** It is critical that the setup is an autonomous triage recommendation system.

# 4 Sequence Diagram

End-to-End Sequence:

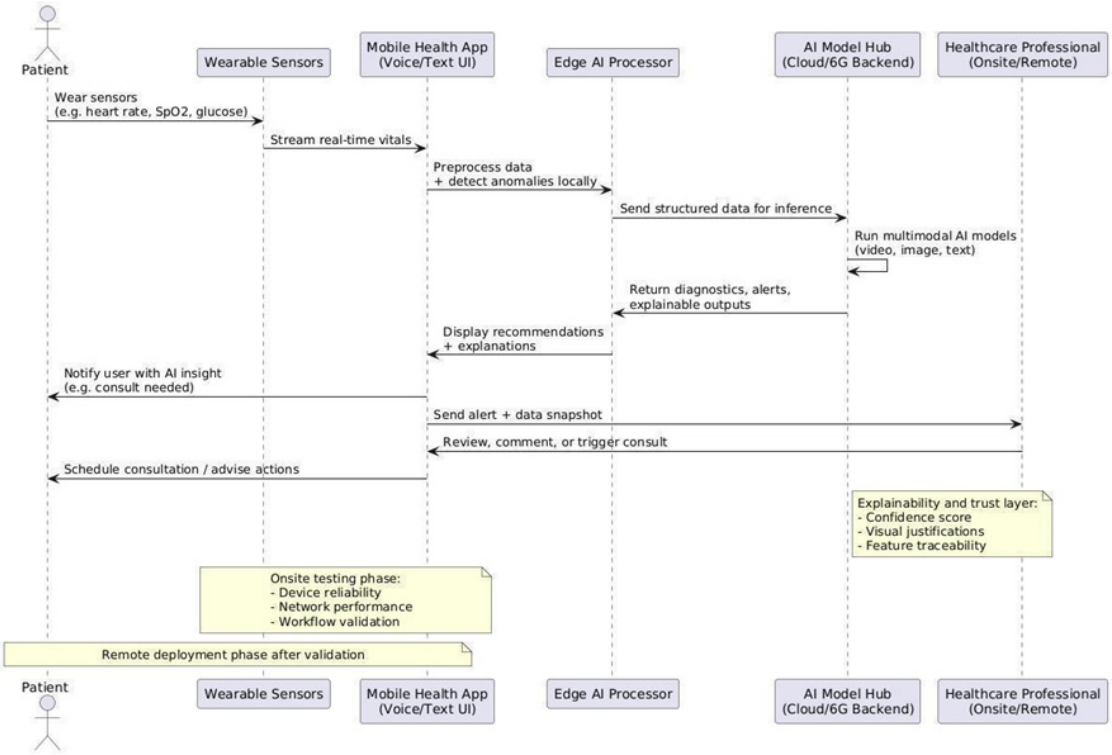



Figure 7.1: Sequence Diagram of our system

Implementation Sequence:

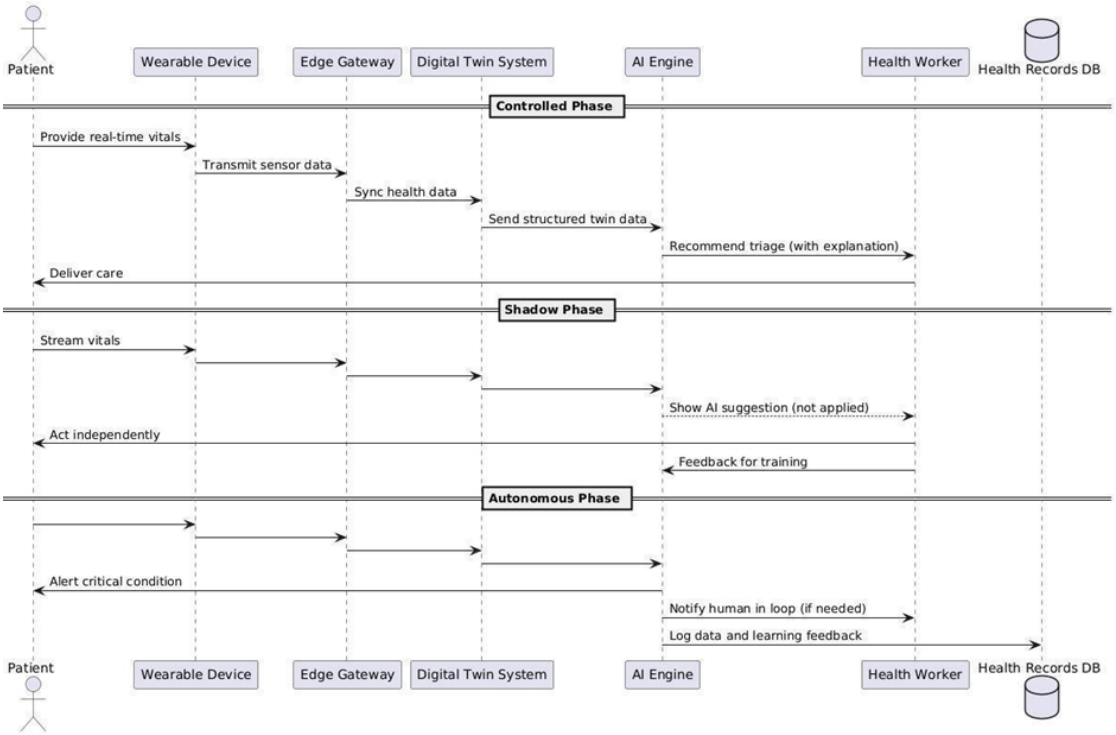


Figure 7.2: Implementational Sequence Diagram of our system

## 5 References

[1] United Nations, Sustainable Development Goals. Available: https://sdgs.un.org/goals.

[2] ITU, AI for Good Initiative. Available: https://aiforgood.itu.int/.

[3] World Health Organization, "Imbalances in rural primary care," 2018. Available: https://apps.who.int/iris/bitstream/handle/10665/346351/WHO-HIS-SDS-2018.58-eng.pdf?sequence=1.

[4] D. Watts et al., "Rural health around the world: challenges and solutions," *Johns Hopkins School of Nursing*, 2024. Available: https://nursing.jhu.edu/wp-content/uploads/2024/06/Rural-health-around-the-world-challenges-and-solutions.pdf.

[5] Tata Elxsi. Available: https://www.tataelxsi.com/.

# Use Case 8: Geospatial Mapping of Otitis Media (OM) and Social Determinants in Remote Northern Territory Communities

**Organization**: NT Health

**Country:** Australia

**Contact person**:

Prasha Sooful Prasha.Sooful@nt.gov.au, +61488478947

## 1 Use case summary Table

| Item | Details |
|---|---|
| Category | Healthcare |
| Problem Addressed | The project addresses the challenge of understanding and tackling otitis media and hearing loss clusters in remote Northern Territory communities by revealing their correlation with specific social determinants like housing, healthcare access, and education, using AI-powered geospatial analysis. |
| Key Aspects of Solution | AI-powered geospatial analysis, high-resolution community data, integration of social and environmental determinants, interactive mapping, data-driven policy recommendations, secure cloud deployment (Azure/ Power BI), advanced statistical modelling (Bayesian hierarchical), and future-proofing with data updates and version control. |
| Technology Keywords | Artificial Intelligence (AI), Geospatial Analysis, Geographic Information Systems (GIS), Bayesian Hierarchical Model, Power BI, Microsoft Azure |
| Data Availability | Public [1] |
| Metadata (Type of Data) | Geocoded Community-Level Data:<br>• Latitude and Longitude Coordinates<br>• Settlement Type (e.g., village, minor town, major town, family outstation)<br>• Mobile Phone Coverage Availability<br>Health Data:<br>• Otitis Media Cases Data<br>• Hearing Loss Cases Data<br>• De-identified Health Records<br>• Diagnostic Assessments (from health clinics)<br>• Service Event Data (for seasonal classification) |

(continued)

| Item | Details |
|---|---|
| Model Training and Fine-Tuning | The core model for this project is a Bayesian Hierarchical Model.<br>Additionally, it will leverage<br>• Power BI's built-in AI features (for initial insights, visualizations) [5].<br>• Python-based machine learning libraries (specifically Pandas and GeoPandas for data handling and spatial operations, which are foundational for applying ML models). |
| Testbeds or Pilot Deployments | The deployment is not a separate phase after development. Implying the pilot deployment is happening alongside development to gather feedback early. |
| Code repositories | Not Available |

# 2 Use Case Description:

## 2.1 Description

This project harnesses the power of AI-driven geospatial analysis to pinpoint and visualise clusters of otitis media and hearing loss within remote communities across Australia's Northern Territory (NT). These clusters are then meticulously overlaid with crucial social determinants such as housing conditions, access to healthcare, and educational attainment, offering a comprehensive view of contributing factors. To ensure unparalleled spatial resolution and granularity, acknowledging the NT's vast and diverse geography, the project utilises geocoded community-level data provided by Telstra. This rich dataset comprises precise latitude and longitude coordinates for 93 distinct NT communities, each enriched with metadata on settlement type (e.g., village, minor town, major town, family outstation) and mobile phone coverage [4]. This granular approach allows for the visualisation of each community as a discrete spatial unit, avoiding reliance on broader, less precise aggregations like postal codes, thereby preserving the fine-grained variations between communities. A custom seasonal classification for the NT has been developed and directly implemented within Power BI using Data Analysis Expressions (DAX) logic. This dynamic classification assigns each service event to a specific season based on the month and health service region. Power BI's time series layers facilitate the filtering and viewing of seasonal trends over time and across different regions, further enhanced by the availability of time sliders. The ultimate aim is to create an interactive geospatial map that provides actionable insights to inform health policy and interventions. The project's core requirements include the creation of an AI-powered, interactive geospatial map that distinctly showcases otitis media clusters within remote Northern Territory communities, overlaid with pertinent socio-economic data. AI technology and analysis are central to generating ear disease heatmaps and spatial clusters, enabling the identification of high-risk areas for otitis media based on social determinants. Datasets concerning hearing loss and ear disease have been acquired through de-identified health records and diagnostic assessments from health clinics. It is a known limitation that this data is manually entered, which introduces a potential margin of error. However, this is mitigated by annual audits conducted by the Continuous Quality Improvement (CQI) manager, nursing team, and data entry officer. Summary public census datasets for each community, specifically from the Australian Bureau of Statistics (ABS) national census in 2021, are incorporated to include key social determinants such as housing conditions, educational attainment, and employment. This data is rigorously overlaid with the 2021 hearing

loss dataset to bolster geospatial analysis and modelling efforts. Acknowledging that these datasets reflect conditions as of 2021, they are considered valid for analysis from 2021 to 2025. This assumption rests on the understanding that significant structural determinants, particularly housing infrastructure in remote communities, are unlikely to undergo substantial change within this timeframe. A known limitation is the ABS's public acknowledgement of persistent undercounting, with a net undercount of Aboriginal and Torres Strait Islander people estimated at approximately 17.4% in the 2021 Post Enumeration Survey [2]. To mitigate potential issues arising from mismatched data vintages across layers, the project will flag the year of origin for each dataset, apply cautious interpretation when comparing data, and commit to updating models and overlays once the 2026 Census data becomes available. Public environment datasets from sources such as [7] and [8] related to water, sanitation conditions, and food security are also integrated. Recognised limitations for these datasets include typically broader geographical levels, necessitating careful consideration of spatial resolution and infrequent updates. To address this, only data corresponding to the specific year of the hearing loss dataset will be used [3]. The project aims to generate data-driven policy recommendations for targeted healthcare interventions in communities grappling with high rates of otitis media. Furthermore, it will produce clear, actionable reports for stakeholders, including policymakers, healthcare providers, and community leaders, highlighting the crucial correlations between social factors and otitis media prevalence.

**Use-case status: The use case is part of a larger research project.**

**Partner(s): Charles Darwin University [6]**

## 2.2 Benefits of use case

Aboriginal and Torres Strait Islander children in Australia have the highest rates of ear disease (otitis media) in the world. It can lead to lifelong impacts making it a critical area for intervention. The Geospatial Mapping of Otitis Media (OM) and its correlation with social determinants in remote Northern Territory communities presents a groundbreaking approach to addressing a critical public health issue for Aboriginal Australians.

This proposal directly targets Good Health and Well-being, as it aims to reduce the health burden of otitis media. By integrating AI-powered geospatial analysis, this project will enable policymakers and health professionals to identify at-risk areas for more targeted service delivery and interventions, effectively improving access and outcomes in underserved regions. Therefore, this use case focuses on addressing health disparities and systemic inequalities faced by Aboriginal communities. By using AI to create heatmaps and spatial clusters, the project will generate actionable insights that can facilitate targeted health interventions. The innovative use of AI and geospatial visualisation holds the potential to evolve into a standard for health policy applications, particularly in remote areas that face similar public health challenges globally. The project offers scalability by integrating up-to-date census data, with future-proofing measures to keep the system evolving in line with changing socio-economic indicators. Future work includes extending this model for broader application across Indigenous populations globally, creating a sustainable and impactful solution to address inequalities in healthcare delivery.

## 2.3 Future works

In future works, we will endeavour to improve data collection, model development, and the creation of new variations or extensions for the same use case, as well as establish reference tools, notebooks, and simulation environments.

# 3 Use case Requirements

- **REQ-1:** It is critical that the system generate an AI-powered, interactive geospatial map visualising otitis media and hearing loss clusters across remote Northern Territory communities, overlaid with relevant socio-economic and environmental data.
- **REQ-2**: It is required that the system shall employ AI technology and geospatial analysis to create ear disease heatmaps and identify distinct spatial clusters of high-risk areas for otitis media based on health records and social determinants.
- **REQ-3**: It is critical that the system integrates and overlays summarised public census datasets (e.g., housing conditions, educational attainment, employment) and public environmental datasets (e.g., water, sanitation, food security) with health data for comprehensive geospatial analysis.
- **REQ-4:** It is critical that the system enable the generation of data-driven recommendations for targeted healthcare interventions in communities identified with high rates of otitis media and hearing loss.

# 4 Sequence Diagram:

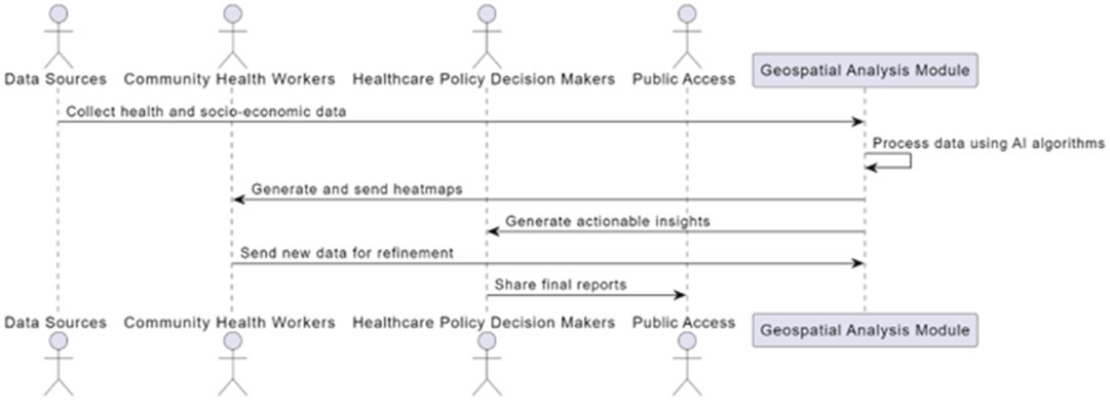



Figure 8.1: Sequence Diagram of our system

# 5 References

[1] AIHW. *Ear and hearing health of Indigenous Australians.* Canberra: Australian Institute of Health and Welfare, 2019. Available at: https://www.aihw.gov.au/reports/indigenous-australians/ear-hearing-health-indigenous-australians

[2] DeLacy, J., Dune, T., & Macdonald, J. *The social determinants of otitis media in Aboriginal children in Australia: Are we addressing the primary causes?* A literature review. BMC Public Health, vol. 19, no. 1, 2019. DOI: 10.1186/s12889-019-6505-7

[3] Jervis-Bardy, J., Carney, A.S., and Duguid, R. *Otitis media in Indigenous Australian children: Review of epidemiology and risk factors.* Journal of Laryngology and Otology, vol. 127, Suppl. 1, pp. S16–S25, 2013. DOI: 10.1017/S0022215112002960

[4] Coleman, L., Thornton, R., Ritter, M. et al. *Microbiological aspects of otitis media in Indigenous Australian children.* Journal of Paediatrics and Child Health, vol. 54, no. 9, pp. 945–951, 2018. DOI: 10.1111/jpc.13997

[5] Microsoft Corporation. *Power BI Dashboard Examples.* Available at: https://powerbi.microsoft.com/en-us/dashboard-examples/

[6] Charles Darwin University. *Official website.* Available at: https://www.cdu.edu.au/

[7] Australian Goverment Department of Climate Change, Energy, the Environment and Water. Available at https://fed.dcceew.gov.au/

[8] Foodbank Available at: https://www.foodbank.org.au/?state=au

# Use Case 9: AI Senior Care Contents: Ensuring Joy and Cognitive Enhancement for Seniors

**Organization**: ESTsoft Inc.

**Country:** South Korea

**Contact Person(s)**:

Kai Gye-Poong Byeon, kai@estsoft.com

Chuck Lee, chuck.lee@estsoft.com

Suy Hwang, suyhwang@estsoft.com

SukJu Kim, sukju@estsoft.com

## 1 Use Case Summary Table

| Item | Details |
|---|---|
| Category | Healthcare |
| Problem Addressed | Our solution aims to address the growing yet under-attended-to emotional and physical challenges faced by seniors (the elderly) in super-aged societies, including deepening social isolation, limited access to digital services, and the widening gap between the supply and demand of caregiving support. |

(continued)

| Item | Details |
|---|---|
| Key Aspects of Solution | 1) AI Senior Care Content<br>• Provides a wide range of structured programs tailored for seniors, delivered through senior centres.<br>• Content genres include AI singing classes (e.g., Korean trot songs), AI physical exercise sessions, and cognitive memory quizzes.<br>• Designed to enhance emotional well-being, physical health, and cognitive stimulation.<br>• ESTsoft offers a solution called AI Studio Perso, a content creation tool that generates AI avatars based on either real individuals (clones) or custom personas. By inputting text, the AI avatar teacher delivers spoken content. Using this solution, ESTsoft produces senior-targeted video content featuring AI avatar instructors [4].<br>• The curriculum and scripts for these videos are collaboratively developed by senior healthcare specialists, university experts in brain education, and certified instructors from the Korea Safety Education Association [2].<br>2) AI Conversational Companion Service<br>• Offers real-time, interactive conversation with AI avatars to reduce social isolation among homebound and single-living seniors.<br>• Integrates with smartwatches to track key health metrics, including sleep quality, step count, and heart rate.<br>• In case of emergencies (e.g., abnormal heart rate or detected falls), the system sends automated alerts to designated caregivers.<br>• Uses image-based food recognition and a health classification engine to provide personalized nutrition analysis and dietary recommendations based on meal photos uploaded by the user.<br>• Delivered in Software Development Kit (SDK) form, the service is compatible with various devices, including TVs, smartphones, tablets, and PCs. |
| Technology Keywords | AI Avatar, Large Language Model (LLM), Speech-to-Text (STT), Text-to-Speech (TTS), Software Development Kit (SDK), Cognitive Training, AI Companion Service, Content Recommendation, User Experience/User Interface (UX/UI) for Seniors |
| Data Availability | Private—collected via deployments in senior centres and partnerships with medical/research institutions |
| Metadata (Type of Data) | text, audio, video, visuals, health metrics (nutrition in calories, sleep time in hours, walk in steps per minute, heartbeat in beats per minute), user interaction logs |
| Model Training and Fine-Tuning | The system leverages models specifically tailored for the senior care domain. A custom-trained LLM is fine-tuned using real data from caregiving scenarios to provide context-sensitive, emotionally attuned conversational responses for both caregiving assistance and companionship.<br>In parallel, STT and TTS models are optimised for elderly speech characteristics, such as slower pacing, pronunciation variation, and lower vocal projection, ensuring reliable voice recognition and natural-sounding responses. These models are further adapted using datasets that reflect linguistic, behavioural, and cognitive patterns commonly observed in older adults. |

(continued)

| Item | Details |
|---|---|
| Testbeds or Pilot Deployments | The solution has undergone extensive field validation through multi-regional pilot deployments, ensuring real-world usability and scalability.<br>• Gimhae City Deployment: AI Senior Care Content has been actively implemented in 36 smart senior centres, with plans to scale the deployment to over 50 facilities nationwide. These centres serve as a core testbed for evaluating content usability, interaction patterns, and health-related engagement metrics.<br>• National Pilot with Ministry of Health and Welfare: A large-scale demonstration involving 500 elderly participants is being conducted under the supervision of the National Rehabilitation Centre. The project tests the effectiveness of the AI Conversational Companion Service across various user profiles and care environments.<br>• International PoC with Japan (SpiralAI) [7]: Collaborative Proof-of-Concept projects are being carried out with SpiralAI and local Japanese governments. These pilots are focused on validating cultural adaptability, multilingual LLM performance and integration into Japan's ageing care ecosystem. |
| Code repositories | Not publicly disclosed |

# 2 Use Case Description

## 2.1 Description

In response to the escalating challenges posed by ageing populations, ESTsoft has developed a comprehensive AI-driven solution to enhance the emotional, cognitive, and physical well-being of seniors [3]. This initiative comprises two core components: AI Senior Care Content and the AI Conversational Companion Service. The content offering includes over 180 specialised programs, such as traditional music, guided exercise, and cognitive training games, currently deployed across senior centres via television-based platforms. A key strength of the service lies in its advanced personalisation capability. Each content module is systematically tagged using a health classification framework adopted from the Ministry of Health and Welfare [5] and the National Rehabilitation Centre [6]. When a senior interacts with the AI avatar, their inputs and behavioural patterns are processed through a classification algorithm that identifies and delivers the most suitable content. As usage data accumulates, ESTsoft aims to evolve this system further using machine learning, ultimately enabling dynamic, data-driven personalisation at scale [1].

Building on this personalised approach, the AI Conversational Companion Service delivers real-time, lifelike interactions through AI avatars. These avatars are equipped with core functionalities including health monitoring, memory assistance, and dementia prevention, all underpinned by a senior-optimised Large Language Model (LLM) and high-precision Speech-to-Text (STT) and Text-to-Speech (TTS) technologies. Real-time feedback is continuously gathered from users across 36 senior centres and partner medical institutions. Integration with wearable devices, such as smartwatches, enables seamless tracking of key health indicators, including heart rate, sleep cycles, and physical activity. These data streams form a robust feedback loop that enhances service quality, facilitates adaptive content delivery, and ensures timely, individualised support for seniors.

**Use Case Status: Ongoing, actively deployed and undergoing continuous development.**

**Partners**: Ministry of Health and Welfare [5], National Rehabilitation Centre, Gimhae City [6], UNIMEWO [14], Eisai Korea [10], SpiralAI (Japan) [7], Yonsei University [8], Seoul National University [13], Dongguk University [9], Konkuk University [11]

## 2.2 Benefits of use case

- Enhance cognitive and physical health through accessible AI content
- Reduce emotional stress and supports social interactions

## 2.3 Future Work

- ESTsoft is committed to advancing its AI-based senior care solutions through a multi-faceted roadmap that combines technological innovation with domain expertise. Central to this effort is the development of a senior-specific LLM that incorporates linguistic, cultural, and contextual patterns unique to elderly users, enabling more intuitive and empathetic interactions. In parallel, the company is enhancing its STT and TTS technologies to more accurately reflect emotional nuance and accommodate the vocal characteristics of older adults. These advancements are supported by an expanding network of Proof-of-Concept (PoC) validations conducted both domestically and internationally, ensuring service effectiveness across diverse user environments. To enable broader adoption, the AI services are being delivered as scalable SDKs, compatible with multiple devices.
- ESTsoft is also reaching out to accelerate its global expansion through strategic partnerships, beginning with active collaborations in Japan and the United States. Underpinning these technological developments is a rigorous content development process carried out in close collaboration with leading institutions: physical exercise modules are co-developed with Yonsei University [13], safety education with the Korea Safety Education Association [12], and cognitive programs with Eisai Korea [11] and Boramae Medical Center at Seoul National University [13]. The validity and impact of this expert-driven approach have been substantiated through academic research and publications in peer-reviewed journals, reinforcing the credibility and real-world applicability of ESTsoft's AI-powered care solutions.

# 3 Use Case Requirements

- **REQ-01**: It is critical that a dedicated LLM is trained on senior caregiving data to ensure context-aware, empathetic, and relevant conversational responses for elderly users.
- **REQ-02**: It is critical that the platform provide a UI/UX specifically optimized for elderly users, focusing on simplicity, clarity, and accessibility to accommodate cognitive and motor limitations.
- **REQ-03**: It is critical that the system utilises highly accurate STT technology adapted to aged voices, accounting for the slower speech, pronunciation variations, and volume differences common among seniors.
- **REQ-04**: It is critical that a personalised content recommendation engine be implemented to tailor educational and engagement materials based on each user's health status and interaction history.
- **REQ-05:** It is critical that the service be compatible across multiple devices, including PCs, TVs, tablets, and mobile phones, via a scalable SDK, ensuring accessibility regardless of hardware.

- **REQ-06**: It is expected that a structured tagging system based on health classifications will be applied to all content, allowing for accurate and consistent matching with user needs.
- **REQ-07:** It is of added advantage that the system integrates with individual health data, enabling feedback loops that support real-time condition monitoring and personalised interventions.
- **REQ-08**: It is of added advantage that an adaptive recommendation algorithm evolves based on ongoing user interactions, enabling dynamic personalisation and increased engagement over time.

## 4 Sequence Diagram

| Step | Request | Processing | Response |
|---|---|---|---|
| (1) User Interaction | The user watches content or interacts with an AI avatar | The app collects user activity and preference data | The app sends collected data to the AI model |
| (2) Preference Transmission | The app sends viewing history and preferences to the AI system | The AI model processes the user's data | The AI prepares for personalized recommendation analysis |
| (3) AI Analysis | The AI system analyzes viewing patterns and user preferences | The system ranks content based on relevance | The AI generates a list of recommended content |
| (4) Recommendation Generation | The AI confirms personalized recommendations | The system organizes and prioritizes the content | The recommendation results are sent to the server |
| (5) Fetching Content | The server searches the database for recommended content | The server prepares content to be displayed | Recommended content is sent to the app |
| (6) Display Recommendation | The app receives recommended content | The app updates the user interface with recommended content | The user checks and selects personalized recommended content |

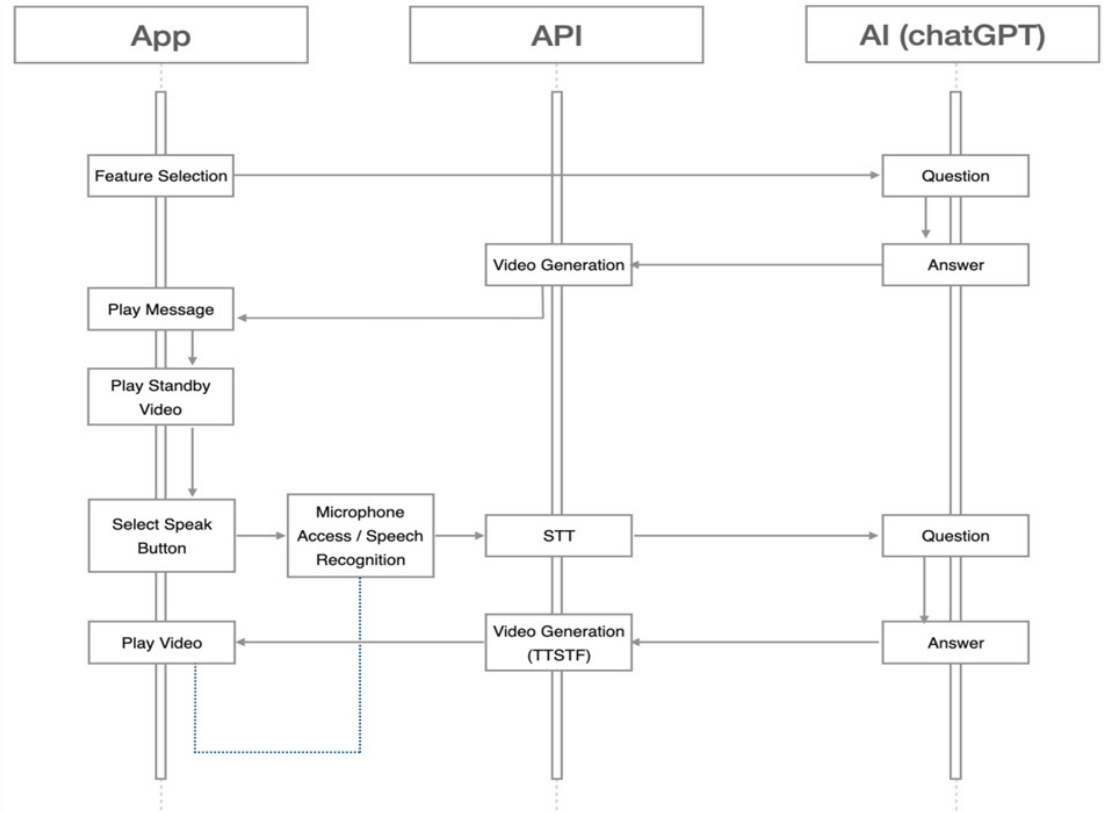



Figure 9.1: System sequence diagram

## 5 References

[1] Korean Journal of Physical Education, Vol. 63, No. 3.

[2] Yonsei University Lifelong Sports Research Laboratory. Available: https://www.yonsei.ac.kr/en_sc/research/lab_intro.jsp.

[3] Eisai Korea & Gachon University. Available: https://www.gachon.ac.kr/sites/eng/index.do.

[4] Senior Care: https://seniorcare.est.ai.

[5] Ministry of Health and Welfare (MOHW), Republic of Korea. Available: https://www.mohw.go.kr/eng/.

[6] National Rehabilitaion Centerl (NRC) of Science & Technology, Republic of Korea. Available: http://www.nrc.go.kr/eng/html/content.do?depth=id&menu_cd=01_02.

[7] go-spiral.ai. Available: https://go-spiral.ai/.

[8] Yonsei University. Available: https://www.yonsei.ac.kr/.

[9] Dongguk University. Available: https://www.dongguk.edu/eng/main.

[10] Eisai Co., Ltd. Available: https://www.eisai.com/company/profile/group/index.html.

[11] Konkuk University. Available: https://english.kku.ac.kr/mbshome/mbs/wwwen/index.do.

[12] Safety Korea (TradeKorea). Available: https://safetykorea.tradekorea.com/.

[13] Seoul National University (SNU). Available: https://en.snu.ac.kr/.

[14] UNIMEWO, https://www.unimewo.com/.

# Use Case 10: AI-mediated Interactive Health Messaging for Community Health Promotion in Low-Middle Income Countries.

**Organization**: American University of Beirut

**Country**: Lebanon

**Contact Person(s)**:

Imad Elhajj – ie05@aub.edu.lb

Aline Germani – ag24@aub.edu.lb

Michelle El Kawak – mk290@aub.edu.lb

## 1 Use Case Summary Table

| Item | Details |
|---|---|
| Category | Healthcare, |
| Problem Addressed | Residents in low- and middle-income countries (LMICs) such as Lebanon often face inequitable access to crucial health information. This disparity is driven by a confluence of factors, including language barriers, low literacy rates, pervasive stigma, and ongoing societal crises. As a result, individuals are often unable to make informed health decisions, leading to poorer health outcomes and exacerbating existing health disparities [1][2]. |
| Key Aspects of Solution | Multilingual chatbot; Speech-to-Text (STT) and Text-to-Speech (TTS) integration; WhatsApp-based deployment; Retrieval Augmented Generation (RAG) with Reasoning + Action (ReAct prompting). |
| Technology Keywords | Large Language Model (LLM), STT, TTS, ReAct prompting, LangChain, Retrieval Augmented Generation (RAG), WhatsApp API, vector databases |
| Data Availability | Publicly sourced health content–World Health Organization (WHO), International Federation of Red Cross and Red Crescent Societies (IFRC), United Nations Children's Fund (UNICEF); Private de-identified interaction logs |
| Metadata (Type of Data) | Text and Audio |

(continued)

| Item | Details |
|---|---|
| Model Training and Fine-Tuning | – LLM Integration: GPT-4 via OpenAI for response generation<br>– Retrieval System: Pinecone vector database for RAG (Retrieval-Augmented Generation)<br>– Hosting & Infrastructure: Managed via DigitalOcean, using Docker for containerization<br>– Speech Capabilities:<br>• Text-to-Speech (TTS) via ElevenLabs—enabling audio responses<br>• Speech-to-Text (STT) via Whisper V3/Replicate—enabling audio input from users<br>– Monitoring & Evaluation:<br>• METABase: custom dashboard for chatbot performance and engagement analytics<br>• Flower: backend tool for real-time technical performance monitoring<br>– Content & Persona Delivery: Synthesia avatars are used to deliver health content in video form |
| Testbeds or Pilot Deployments | Active, with ongoing deployment and iterative refinement. (Lebanon and Jordan) |
| Code repositories | Not Available |

# 2 Use Case Description

## 2.1 Description

Health misinformation, language barriers, and inequitable access to trustworthy health information continue to threaten community well-being in low- and middle-income countries (LMICs), especially among displaced and vulnerable populations [1]. These barriers are often exacerbated by low literacy levels, social stigma around health topics, and a lack of culturally sensitive communication platforms. To address these issues, HIBA (Health Information Bot Assistant) was developed as a WhatsApp-based, AI-powered chatbot that delivers interactive, multilingual health messaging in both text and audio formats [9][10]. Its key objective is to improve equitable access to validated health information by leveraging AI tools such as LLMs, STT, and TTS to enhance inclusivity and user engagement. HIBA's solution architecture is based on a RAG framework integrated with ReAct prompting, enabling the chatbot to deliver accurate, relevant, and user-friendly responses while preventing hallucinations [2][3].

HIBA has been piloted in Lebanon and Jordan with refugee and host communities, integrating feedback from multi-phase user testing to optimize content flow, accessibility, and tone. The chatbot supports 28 languages, including Arabic dialects, and applies community-validated ethical guardrails to ensure that content is informational and not diagnostic. The system architecture incorporates Pinecone for vector search, ElevenLabs for human-like TTS, and GPT-4 for language understanding. For STT, the open-source model Whisper V3 from OpenAI was utilized, which worked well with multiple Arabic dialects, in conjunction with an added layer of LLM inference that tries to match the utterances to intents. As part of the HIBA development process, a comprehensive four-level testing framework was implemented to ensure the quality,

usability, and effectiveness of the chatbot [4][5]. The goal was to continuously improve the system by integrating feedback from internal teams, target users, and broader deployment settings. The initial testing layer, considered a quality control test, focused on generating a high-quality multilingual corpus derived from health topics identified through FGDs. This database was enriched with data from trusted sources like WHO, IFRC, and UNICEF. The curated corpus of 165 common health queries was stored in a vector database and tested for accuracy, clarity, and cultural relevance. Responses were refined to remove technical jargon, US-centric data, and any diagnostic content. Subsequently, two phases of user testing were conducted: (1) A workshop with 6 community members tested the Telegram interface, conversational flow, and language handling. Feedback was gathered via FGDs. (2) 20 participants from Lebanon (CHWs and LRC/IFRC volunteers) tested the WhatsApp version over five days [6]. Feedback was collected using the METABase dashboard and reflective journaling. The third layer of testing included 50 participants from vulnerable communities in Lebanon and Jordan who tested HIBA for two weeks. Post-trial assessments via phone interviews and pre-defined metrics provided both qualitative and quantitative feedback on accessibility, clarity, and user trust. Finally, a social marketing campaign deployed HIBA in a Lebanese city with a diverse urban setting. Materials tailored to subgroups (e.g., students, parents) were disseminated via NGOs, schools, and parishes using QR-coded posters, flyers, and videos. Recently, the campaign extended to displacement centers housing war-affected populations. The campaign's reach and impact were reinforced by a community mobilizer and evaluated through performance monitoring on the chatbot dashboard. Overall, the chatbot responses were assessed using a structured rubric and Likert scale, evaluating clarity, appropriateness, and trustworthiness. Commonly asked topics and emergent patterns were also tracked. The expected impact includes increased health literacy, reduced misinformation, and improved community engagement in health decision-making, especially for populations that typically fall outside the reach of conventional health promotion strategies [11][13][14].

**Use Case Status: Active, with ongoing deployment and iterative refinement.**

**Partners:**

- Lebanese Red Cross (LRC)—Youth Sector [15],
- Jordan National Red Crescent Society (JNRCS) [16]
- IFRC MENA Regional Office [17]

## 2.2 Benefits of use case

HIBA ensures equitable access to health information, particularly for marginalized populations. By providing accurate, evidence-based health messaging, the chatbot empowers users with preventive healthcare knowledge, reducing reliance on overburdened healthcare systems. It also facilitates early identification of health concerns, enabling timely medical interventions.

HIBA functions as an AI-driven educational tool for health literacy. By using interactive Q&A, multilingual content, and speech-based engagement, it ensures that people with low literacy levels or disabilities can access relevant, understandable health information. Additionally, it supports continuous learning for community health workers and volunteers.

By targeting underserved populations in LMICs, HIBA reduces health disparities caused by socioeconomic and linguistic barriers. The AI-driven chatbot removes stigma, enabling individuals - especially women and refugees - to access trusted health guidance anonymously

and privately. Furthermore, HIBA's AI-driven insights help public health authorities identify gaps in health literacy and emerging health trends, informing more inclusive healthcare policies.

## 2.3 Future Work

Future development of HIBA will focus on two key areas: enhanced AI model development and new variations for broader impact.

- Enhanced AI Capabilities for Nuanced Understanding: We'll significantly improve HIBA's AI, focusing on natural language processing (NLP) to accurately understand regional dialects and complex medical queries in Arabic and other underrepresented languages. This includes integrating sentiment analysis and emotional AI to detect user distress and adapt responses for better mental health support, alongside implementing adaptive learning for continuous self-improvement in accuracy and relevance [12].
- Broader Impact through Localized Deployment: Our goal is to create localized HIBA versions tailored to specific linguistic and cultural groups (e.g., Kurdish, Sudanese Arabic) with relevant health messaging. This will enable deployment in humanitarian and crisis settings like refugee camps, where health information is critical and access to professional care is limited.
- Strategic Growth through Technology and Collaboration: We plan to upgrade our technology infrastructure with cloud solutions and real-time analytics for robust performance. This will be supported by ongoing data collection and dataset creation of validated, localized health content. Crucially, we'll actively seek collaborations with public health authorities, humanitarian organizations, and AI research labs to co-develop features, share resources, and expand our pilot programs globally.

# 3 Use Case Requirements

- **REQ-01:** It is critical that a multilingual STT/TTS engine (Arabic, English, French + 25 other languages supported) and a geographically determined multilingual conversational flow are incorporated into the system.
- **REQ-02:** It is critical that ethical guidelines be followed to prevent diagnostic/treatment guidance
- **REQ-03:** It is expected that there is a General Data Protection Regulation (GDPR)-compliant infrastructure and WhatsApp API integration
- **REQ-04:** It is critical that there is a Retrieval-Augmented Generation (RAG) framework with a vector database integrated
- **REQ-05:** It is of added advantage that a real-time dashboard for performance monitoring is integrated.
- **REQ-06:** It is expected that a culturally and gender-sensitive voice and avatar design is integrated into the interface.

## 4 Sequence Diagram

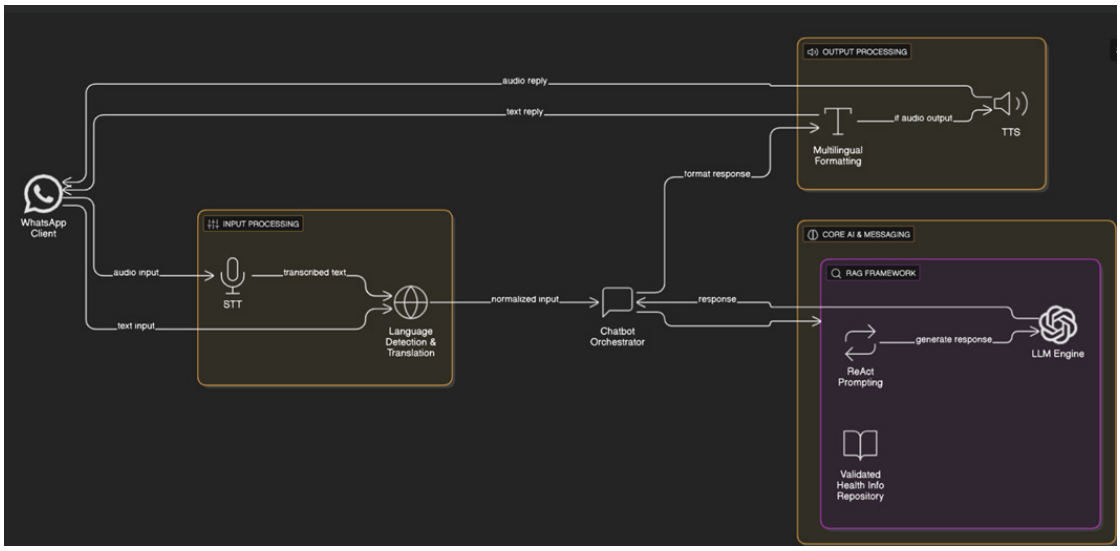



Figure 10.1: Sequence Diagram of our system

## 5 References

[1] Baclic, O., *et al., Artificial intelligence in public health: Challenges and opportunities for public health made possible by advances in natural language processing*. Canada Communicable Disease Report, 2020. 46(6): p. 161.

[2] Ciecierski-Holmes, T., *et al., Artificial intelligence for strengthening healthcare systems in low-and middle-income countries: a systematic scoping review*. npj Digital Medicine, 2022. 5(1): p. 162.

[3] Bhirud, N., *et al., A literature review on chatbots in healthcare domain*. International journal of scientific & technology research, 2019. 8(7): p. 225-231.

[4] Wang, X., *et al., ChatGPT: promise and challenges for deployment in low-and middle-income countries*. The Lancet Regional Health–Western Pacific, 2023. 41.

[5] Walwema, J., *The WHO health alert: communicating a global pandemic with WhatsApp*. Journal of Business and Technical Communication, 2021. 35(1): p. 35-40.

[6] Harris, J., *et al., Evaluating Large Language Models for Public Health Classification and Extraction Tasks*. arXiv preprint arXiv:2405.14766, 2024.

[7] Zhao, W.X., *et al., A survey of large language models*. arXiv preprint arXiv:2303.18223, 2023.

[8] Chang, Y., *et al., A survey on evaluation of large language models*. ACM Transactions on Intelligent Systems and Technology, 2024. 15(3): p. 1-45.

[9] Battineni, G., N. Chintalapudi, and F. Amenta. *AI chatbot design during an epidemic like the novel coronavirus*. in Healthcare. 2020. MDPI.

[10] Aharoni, E., *et al., Attributions toward artificial agents in a modified Moral Turing Test*. Scientific Reports, 2024. 14(1): p. 8458.

[11] Gams, M. and S. Kramar, *Evaluating ChatGPT's consciousness and its capability to pass the Turing test*: A comprehensive analysis. Journal of Computer and Communications, 2024. 12(03): p. 219-237.

[12] Fraser, H., *et al., Comparison of diagnostic and triage accuracy of Ada health and WebMD symptom checkers, ChatGPT, and physicians for patients in an emergency department: clinical data analysis study*. JMIR mHealth and uHealth, 2023. 11(1): p. e49995.

[13] Patel, A.B., *et al., M-SAKHI—Mobile health solutions to help community providers promote maternal and infant nutrition and health using a community-based cluster randomized controlled trial in rural India: A study protocol*. Maternal & Child Nutrition, 2019. 15(4): p. e12850.

[14] Luxton, D.D., S.L. Anderson, and M. Anderson, *Ethical Issues and Artificial Intelligence Technologies in Behavioral and Mental Health Care, in Artificial Intelligence in Behavioral and Mental Health Care*. 2015, Elsevier Academic Press. p. 255-276.

[15] *Lebanese Red Cross - Youth Sector*. 2025; Available from: https://www.redcross.org.lb/.

[16] *IFRC National Societies Directory Jordan National Red Crescent Society*. 2025; Available from: https://www.ifrc.org/national-societies-directory/jordan-national-red-crescent-society.

[17] *International Federation of Red Cross and Red Crescent Societies (IFRC) Middle East and North Africa (MENA) Regional Office*. 2025; Available from: https://www.ifrc.org/who-we-are/people-and-structures/global-leadership-team/regional-director-mena.

# Use Case 11: Necrosis Detection in Meningioma Images

**Organization**: International Institute of Information Technology (IIIT) Bangalore

**Country**: India

**Contact Person(s)**:

Hitesha Mukherjee hitesha144@gmail.com

TK Srikanth tk.srikanth@iiitb.ac.in

Ramesh Kestur rameshkestur@iiitb.ac.in

Anita Mahadevan mahadevananita@gmail.com

Nandeesh BN bnnandeesh3@gmail.com

Shilpa Rao shilpagk.rao@gmail.com

## 1 Use Case Summary Table

| Item | Details |
|---|---|
| Category | Healthcare |
| Problem to be addressed | Addressing the need for a robust Artificial Intelligence (AI) pipeline that would be useful in the detection of Brain Tumours to reduce the workload of pathologists |
| Key aspects of solution | Machine Learning (ML) and Ensemble Pipelines in Necrosis Detection in Meningioma Histopathological Images |
| Technology keywords | Necrosis, Medical Image Analysis, Histopathogical Images, Meningioma, Brain Cancer, AI for Cancer, Machine Learning for Healthcare |
| Data Availability | Private |
| Metadata (type of data) | Image Data |
| Model Training and Fine-Tuning | Created a pipeline of AI Algorithms |
| Testbeds or Pilot Deployments | Hosted on University Servers |
| Code repositories | Private repositories |

# 2 Use Case Description

## 2.1 Description

### Context and Background

Brain cancer is a grave malignancy type that is a consequence of unconstrained and unrestricted multiplication of cells in the brain, generally engendered by a malignant brain tumour.

Meningiomas are among the most common tumours of the Central Nervous System (CNS). These sorts of tumours, which have a varied growth potential, emerge from the meninges, the membranes that coat the brain and spinal cord, and generally do not metastasize.

**Necrosis** - It is a dead region, and there is an absence of nuclei. It consists of small, dust-like structures. The cells are broken down along with cell borders, and there is a larger quantity of cell debris. There are many cells and connective tissues are fewer. Texture changes occur slowly in necrosis.

A biopsy involves examining a tissue under a microscope to determine the presence or extent of disease spread.

Histopathological (biopsy) analysis is a painstaking, prolonged, and labour-intensive skilled job, dependent on the expertise of the pathologist. With the availability of high-resolution digital images of slides, referred to as Whole Slide Images (WSI), digital histopathology opens the potential to automate and improve parts of this analysis process, with the goal of more efficient and less error-prone diagnosis, and reducing the workload of pathologists.

Analysis of histopathological (biopsy) slides by pathologists is considered the gold standard in determining the presence and nature of diseases such as cancer. Thin sections of tissue extracted during biopsies are fixed on a glass slide and stained with specific chemicals, e.g., Haematoxylin and Eosin (H&E) stains, and then observed under a microscope at different levels of magnification.

Different magnification levels are used for different types of analysis, from which are derived the final impressions and diagnosis, such as categorization of tumours as malignant or benign, and if cancerous, the grade of the disease. In our study, the selected glass slides, stained with H&E, were scanned using a Leica SCN400 microscopic whole-slide scanner with nominally 20x magnification.

### Aims and Objectives

It was observed that there is a lack of publicly accessible labelled datasets of meningioma whole slide images. Hence, the first objective in the current work was the generation of a sizable, high-quality, expert validated, ground truth dataset of pathology features of meningioma whole slide images, which could be used for this research, and would also form the basis of future work on neuropathology analysis.

The second and primary objective was the development of ML pipelines for detecting regions of necrosis in whole slide images of meningioma. This includes a comparison of the performance of various ML and ensemble algorithms, as well as the evaluation of performance metrics such as specificity and sensitivity of such algorithms with varying sizes of necrotic regions.

## Technological Approach

The proposed approach includes building a pipeline consisting of data acquisition, data pre-processing, data exploration, feature selection, model building, and data visualization. These pipelines are seamlessly extensible.

During the pre-processing stage, the WSI, along with the ground truth annotation indicating the locations of regions, is taken by the patch-based framework. In this approach, small positive and negative patches are extracted from the set of training WSIs. Local mini-patches (tiles) around 256 × 256 are sampled from the large WSI around the annotated regions. Tiles are classified according to their biological feature, such as necrosis, mitosis, etc., relative to the annotated polygons of the WSI.

The retrieved tile is turned into a 256 × 256 ×3 Numpy array. Tiles with predominantly whitespace and small artifacts have been removed.

OpenSlide manages the properties and characteristics of distinct vendor formats at run-time between the OpenSlide API and a particular vendor's data format.

We use a thresholding process to segregate the tissue from the background whitespaces, which are present in the whole slide images. Since white tiles are known artifacts that do not have any useful features, those tiles can be eliminated.

The threshold that we have set is 5% of tissue area based on the literature survey. If a tile has less than 5% tissue by area, implying 95% or more whitespace, then that tile is eliminated. Handcrafted features, relating to colour distribution, textures, and morphology, are extracted using Gray-Level Cooccurrence Matrix (GLCM), histogram, and region-based feature measures.

The corresponding features are fused together in different combinations, and a number of machine learning and boosting algorithms are applied to get the desired model. The effect of stain normalization, augmentation, and Principal Component Analysis (PCA) are also studied.

The following selected features were utilized for building the ML and Ensemble pipeline: Histogram, GLCM Features, Region Features like eccentricity, maximum area, extent, and perimeter. Grid search CV, along with 5-fold cross-validation, was used for tuning the hyperparameters.

A combination of qualitative and quantitative statistical measures is used to compare the classification performance of a number of these algorithms to determine the best models that can augment the clinical decision-making process of the pathologists. The performance of the algorithms across a set of whole slide images, as well as the performance across tiles with varying amounts of necrosis regions, is also studied.

Adaboost has given optimum performance among the Boosting algorithms. Logistic Regression has given optimum performance among the Classical Algorithms.

## Partners

Neuropathology, a relatively niche specialization of pathology, involves analysis of biopsy tissues of the brain, spine and nerves. Working with the Neuropathology Lab at National Institute of Mental Health and Neurosciences (NIMHANS), Bangalore, we explore the use of computational

tools, and specifically machine learning algorithms, in the analysis of histopathology images along with E-HRC lab at the International Institute of Information Technology – Bangalore (IIITB)

## 2.2 Benefits of the use case

Our use case pertains to this pertinent goal of good health and well-being.

## 2.3 Future Work

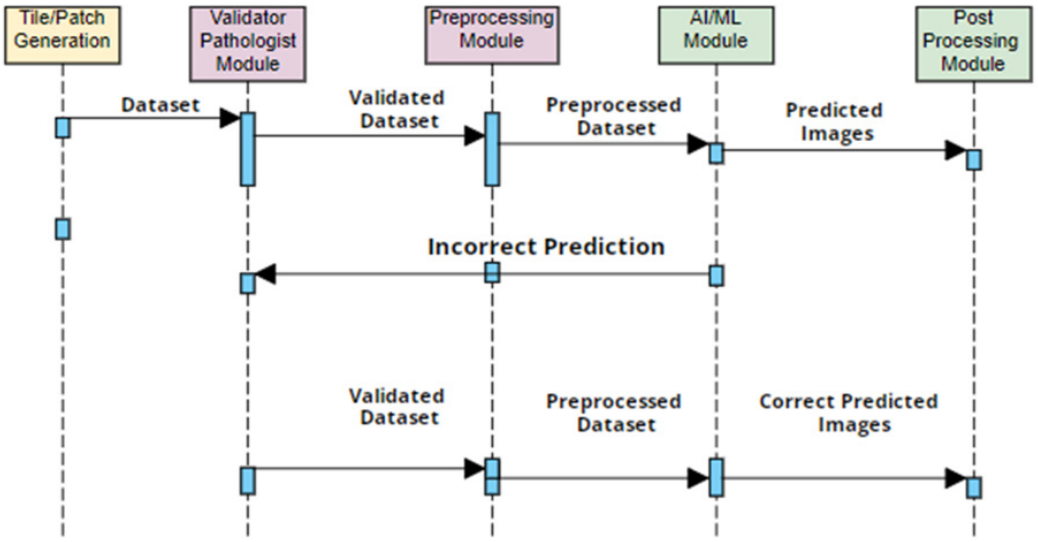
This work was carried out under the constraints of limited data sets and limited class labels. These constraints were addressed by formulating necrosis detection as an ML problem. We set up a complete pipeline that is seamlessly extensible. The results are encouraging, and the work demonstrates the potential of AI/ML in Computational Neuropathology for screening and detection of malignancies of tumours of the CNS.

Annotated data in digital histopathology is hard to acquire, unlike problems such as face detection, which is available commonly, since analysing and labelling histopathology data requires the involvement of expert pathologists. The availability of larger datasets will enable further validation of the approach and enable incorporation into a tool that can be used by pathologists. These will need to ensure adequate performance in the analysis of WSI and ensure high rates of specificity and sensitivity across demographics and scanners, with repeatability of results. A methodology must be defined for validation of the algorithm on a much larger set of WSI samples, covering multiple scanners and patient demographics.

# 3 Use Case Requirements

- **REQ-1:** Need to have annotated histopathological images verified by a committee of pathologists before proceeding to build the use case pipeline.
- **REQ-2:** All the artifacts, whitespace, noise etc must be removed before further processing.
- **REQ-3:** Augmentation needs to be performed depending on the dataset size.
- **REQ-4:** Handcrafting of features must be done before the Machine Learning and Ensemble Algorithms are applied.

## 4 Sequence Diagram




## 5 References

[1] H. Mukherjee et al., "Necrosis Detection in Meningioma Images," presented at the IEEE SPS Winter School on Biomedical Signal & Image Processing, RIT Bangalore, 2019.

[2] M. N. Gurcan, L. E. Boucheron, A. Can, A. Madabhushi, N. M. Rajpoot, and B. Yener, "Histopathological Image Analysis: A Review," IEEE Rev. Biomed. Eng., vol. 2, pp. 147–171, 2009, doi: 10.1109/RBME.2009.2034865. Available: http://ieeexplore.ieee.org/document/5299287/.

[3] H. Sharma et al., "Appearance-based necrosis detection using textural features and SVM with discriminative thresholding in histopathological whole slide images," in 2015 IEEE 15th International Conference on Bioinformatics and Bioengineering (BIBE), Belgrade, Serbia: IEEE, 2015, pp. 1–6. doi: 10.1109/BIBE.2015.7367702. Available: http://ieeexplore.ieee.org/document/7367702/.

[4] A. T. Hale, D. P. Stonko, L. Wang, M. K. Strother, and L. B. Chambless, "Machine learning analyses can differentiate meningioma grade by features on magnetic resonance imaging," *Neurosurg Focus*, vol. 45, no. 5, p. E4, Nov. 2018, doi: 10.3171/2018.8.FOCUS18191

[5] K. Kourou, K. P. Exarchos, C. Papaloukas, P. Sakaloglou, T. Exarchos, and D. I. Fotiadis, "Applied machine learning in cancer research: A systematic review for patient diagnosis, classification and prognosis," Computational and Structural Biotechnology Journal, vol. 19, pp. 5546–5555, Jan. 2021, doi: 10.1016/j.csbj.2021.10.006. Available: https://www.sciencedirect.com/science/article/pii/S2001037021004281.

# Use Case 12: AI-Based Horse Identification and Gait Analysis.

**Organization**: AIforPet

**Country**: South Korea

**Contact Person(s)**:

Euna Huh, eahuh@aiforpet.com

Seonghoon Kim, shkim@aiforpet.com

Jihyun Seo, jhseo@aiforpet.com

Kiyoung Jang, ky.jang@kra.co.kr

## 1 Use Case Summary Table

| Item | Details |
|---|---|
| Category | Healthcare |
| Problem Addressed | Horses, known for their close bond with humans, are frequently exposed to health risks such as lameness, which in severe cases can lead to the loss of the animal. Horse owners strive to minimize these risks and seek ways to detect problems as early as possible. At the same time, veterinarians are in need of additional tools that can enhance diagnostic objectivity and support more systematic and proactive health management. |
| Key Aspects of Solution | Image segmentation, Pose Estimation, Gait Analysis, Image Processing, Mobile apps |
| Technology Keywords | Horse Identification, Horse Lameness Detection, Animal Healthcare |
| Data Availability | Private; Internal data used for model training and validation |
| Metadata (Type of Data) | Image data, Video data, Annotation (Segmentation, Pose Estimation) data |
| Model Training and Fine-Tuning | • Real-time Multi-Person Pose Estimation (RTMPose) for pose estimation<br>• Segmentation Transformer (SegFormer) and Mask Region-Based Convolutional Neural Network (Mask R-CNN) for horse passport generation<br>• Motion Contrastive Anomaly Detection (MoCoDAD) and Convolutional Neural Network–Long Short-Term Memory (CNN-LSTM) models for lameness detection |
| Testbeds or Pilot Deployments | The system has been evaluated in real-world settings through pilot deployments, demonstrating its practical feasibility and user acceptance [1]. |
| Code repositories | Private |

# 2 Use Case Description

## 2.1 Description

The AI-powered equine health management platform (TT-care Equine) [1] is designed to solve two critical problems in equine welfare: horse identification and early detection of lameness. Horses, which have long had a close relationship with humans, are vulnerable to risks such as lameness [2], and misidentification and mismanagement are critical to the preservation of the species. Traditional methods to address these issues rely heavily on manual observation and subjective assessment, resulting in inefficient, inconsistent, and delayed interventions.

The proposed platform integrates a smartphone-based platform with advanced AI algorithms to provide two key features. First, the horse identification module uses a common smartphone camera to capture unique physical patterns in images of the horse's face, body, and legs. AI models automatically extract the unique features and compare them to a digital passport system to verify the horse's identity. Digitizing the horse registration process increases data accuracy, reduces administrative burden, and improves accessibility for horse owners and veterinarians.

Second, the lameness detection module analyzes gait patterns by applying AI-based pose estimation techniques from video footage taken with a smartphone. The system detects subtle movement abnormalities that may indicate the early stages of lameness, providing objective data to support a veterinary diagnosis. Unlike traditional methods that require specialized equipment such as force meters, treadmills, and inertial measurement units (IMUs), this approach uses widely available mobile devices, significantly lowering the barrier to regular health monitoring. The platform also facilitates preventive health care strategies by enabling longitudinal tracking of gait data.

The proposed solution addresses key limitations of existing practices, including reliance on subjective judgment [3], high cost of measurement equipment, and lack of continuous data collection and analysis systems. It also contributes to animal welfare and sustainable livestock production by facilitating ongoing and proactive equine health management.

**Use Case Status: Pilot**

**Partners: Korea Racing Authority (KRA)**

## 2.2 Benefits of use case

This attempt to combine horse management and AI has the potential to improve the welfare of horses through ongoing population protection and health management, while also contributing to sustainable livestock farming.

## 2.3 Future Work

Building on the current platform, AIforPet will focus on three key areas to further advance animal welfare through AI-powered solutions.

First, the scale and diversity of datasets will be expanded by collecting images and gait information from a broader range of horse breeds, age groups, and activity types. This

enhancement is expected to improve the robustness and generalizability of the models across different environments and conditions.

Second, the AI models will be further developed by integrating multimodal data, such as combining gait analysis with vital signs or behavioral cues, enabling faster and more accurate detection of health abnormalities. This approach not only aims to improve diagnostic precision but also to extend the platform's applicability to species beyond horses.

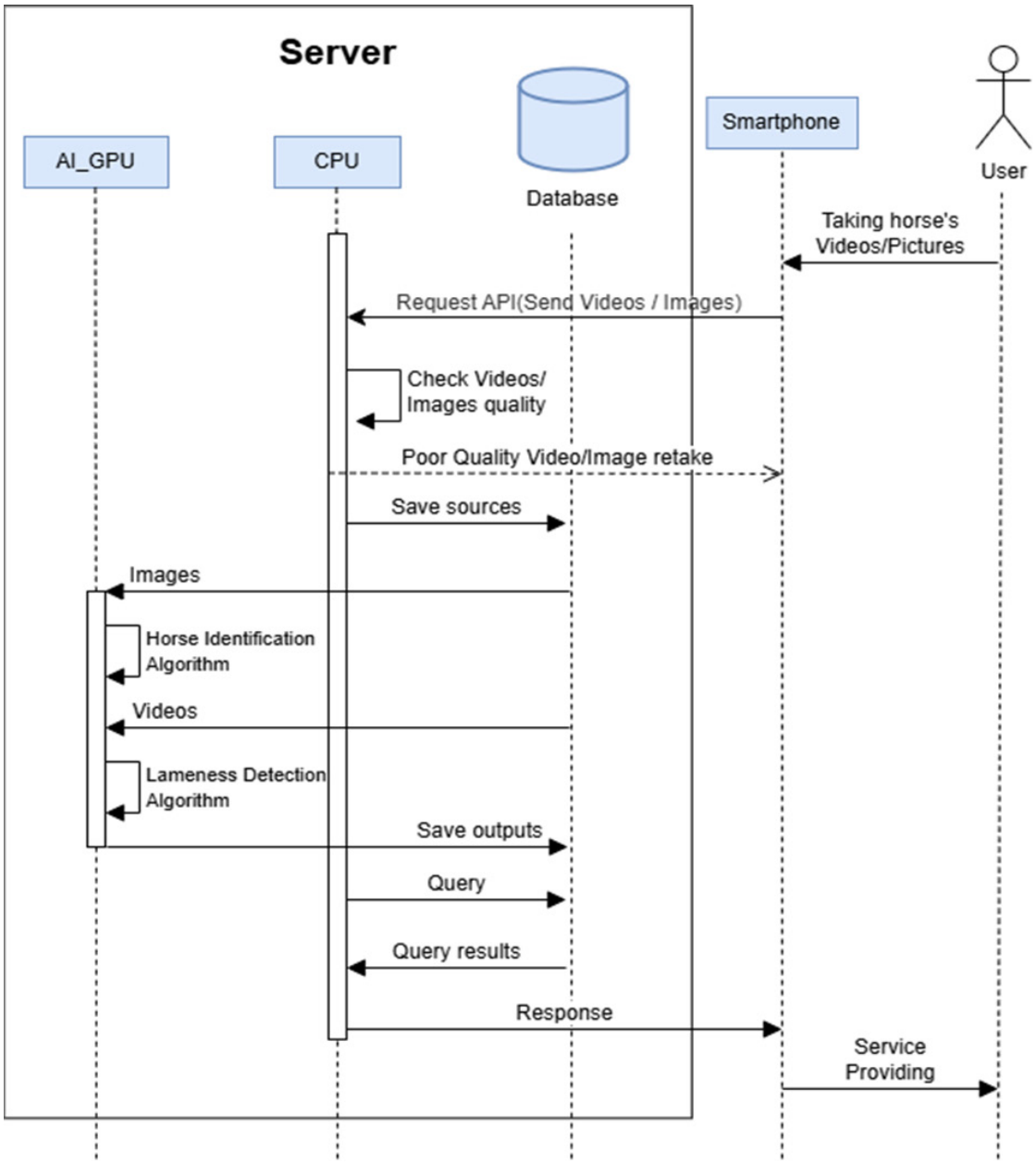
Third, a comprehensive animal welfare monitoring system will be established to support continuous health tracking and disease prevention across multiple animal species [4]. Collaboration with veterinary institutions, farms, and animal welfare organizations will be pursued to create a sustainable ecosystem in which AI technology contributes to proactive animal care and protection.

Ultimately, through continuous innovation and strategic partnerships, AIforPet seeks to realize its vision of enhancing the welfare of diverse animal species and promoting harmonious coexistence between humans and animals.

## 3 Use Case Requirements

- **REQ-01:** The system should collect standardized images and videos to ensure consistent, high-quality input for AI inference. The quality and accuracy of the input data directly affects the performance of the horse identification and lameness detection models, so all data should clearly capture relevant parts of the horse.
- **REQ-02:** The AI systems are meant to help and support veterinarians' clinical judgment, not replace it. Diagnostic results should be used as reference information, and the final decision should be left to the veterinarian.
- **REQ-03:** The system should be able to perform horse identification and lameness detection simultaneously or separately, depending on user requirements. Operational flexibility is essential to accommodate different use cases in a field environment.
- **REQ-04:** AI models must be continuously updated with newly collected data to improve recognition accuracy for horse identification and early detection of lameness. Continuous learning is critical to adapt to changes in breed, environment, and health conditions.
- **REQ-05:** The system should include mechanisms to encourage long-term monitoring by notifying users when regular checks or new data collection is recommended. This feature supports sustainable animal welfare practices by facilitating regular health assessments.

## 4 Sequence Diagram




## 5 References

[1] AI FOR PET, "TTcare Equine: AI-based Equine Health and Lameness Detection," *AI for Pet*, 2024. [Online]. Available: https://www.aiforpet.com/ttcare-equine/. [Accessed: 6-Jun-2025].

[2] Stashak, T. S. 2002. Examination for lameness. pp. 113–183. In: Adams' Lameness in Horses, 5th eds. ed., Lippincott Williams & Wilkins, Philadelphia.

[3] Kalinkina, Galina, et al. 2024. "Digital transformation of methods for assessing the horse exterior characteristics." BIO Web of Conferences. Vol. 108. EDP Sciences.

[4] A. I. Forpet Co., Ltd., "Service providing system for predicting and diagnosing joint condition and method thereof," Korean Patent KR1020220006399, filed Jan. 17, 2022.

# Use Case 13: AI-Driven Optimization of Clinical Trial Preparation: From Protocol Design to CRO Selection

**Organization:** MEDIAIPLUS Inc.

**Country:** Republic of Korea

**Contact Person(s)**:

Primary contact: Nayeon Ryoo, biz@mediaiplus.com

Secondary contact: Ji Hee Jung

## 1 Use Case Summary Table

| Item | Details |
|---|---|
| Category | Healthcare |
| Problem Addressed | Fragmented manual workflows in clinical trial design and Contract Research Organization (CRO) selection lead to inefficiencies, high costs and delays |
| Key Aspects of Solution | • Improve the speed and accuracy of clinical trial preparation<br>• Optimize study design using LLM insights (MediC)<br>• Automate CRO matching with RAG (FiCRO) |
| Technology Keywords | Large Language Models (LLM), Retrieval-Augmented Generation (RAG), Natural Language Processing (NLP), Clinical Trial Optimization, AI Partner Matching |
| Data Availability | More than 720,000 structured and unstructured datasets from global clinical trial registries, 29,500 CRO profiles, and biomedical research outputs |
| Metadata (Type of Data) | Structured tabular data on clinical trials and CROs, including protocols, phases, outcomes, regions and sponsor/CRO attributes. |
| Model Training and Fine-Tuning | Custom LLMs fine-tuned on biomedical and clinical corpora; RAG models enhanced using real-world trial matching data and continuous user feedback loops |
| Testbeds or Pilot Deployments | Ongoing pilot deployments with biotech startups and clinical research institutions in Korea; planned expansion with multinational CROs and pharma sponsors |
| Code repositories | Not publicly available due to data sensitivity and commercial confidentiality. |

# 2 Use Case Description

## 2.1 Description

Clinical trials often suffer from inefficiencies due to manual processes and fragmented CRO selection. MEDIAIPLUS addresses this by offering AI-based solutions to automate trial design and optimize CRO matching, reducing delays and improving outcomes. Its two key engines (MediC and FiCRO) provide advanced inferences such as protocol recommendations, recruitment analysis, and CRO suggestions. These capabilities support real-time decision-making across feasibility analysis, sponsor–CRO matching, and trend identification. As a result, the platform helps shorten planning cycles, lower costs, and improve trial alignment and success rates.

**Use Case Status:** Pilot deployments are currently underway with multiple biotech and medical device companies. Blisstech serves as a key validation partner under the AX Project, while discussions with other international CROs and sponsors are ongoing.

**Partners:** Biotech, pharmaceutical company, medical device company, CROs

## 2.2 Benefits of use case

By accelerating clinical trial design and reducing costs, MEDIAIPLUS improves timely access to innovative treatments and supports R&D efficiency in therapeutics. nearly-stage deployments, MEDIAIPLUS improved clinical trial preparation by standardizing protocol design through MediC and automating CRO matching with FiCRO. These combined efforts increased overall trial success rates by 25%, and reduced the time required to collect and align clinical trial data by 85%—directly supporting faster patient access to critical therapies. MEDIAIPLUS enhances biotech R&D capacity by integrating AI into both protocol design and CRO operations.Through its scalable AI platforms—MediC for trial protocol structuring and FiCRO for CRO matching—the company has enabled automation in clinical planning workflows. This innovation has resulted in a 10% reduction in trial outsourcing costs by matching sponsors with CROs that best align with their therapeutic and regional needs, strengthening the technological infrastructure of trial preparation.

MEDIAIPLUS fosters cross-border collaboration in clinical innovation by connecting trial sponsors, CROs, and research institutions through a unified AI-driven ecosystem. The platform serves as a technical enabler for interoperable workflows, data integration, and decision support across regions. Ongoing engagements with biotech startups and research hubs in Korea and Southeast Asia highlight its role in scaling AI applications for clinical research in both emerging and developed markets.

## 2.3 Future Work

We are advancing our platform with key enhancements focused on automation, precision, and integration. A major initiative includes leveraging large language models (LLMs) to automate protocol generation, streamlining trial design and reducing manual effort. The CRO scoring system will also be upgraded with predictive models for more accurate, data-driven evaluations. To support these developments, FiCRO requires access to broader clinical trial datasets, scalable cloud infrastructure, and improved AI explainability tools to ensure transparency and user trust. Expansion efforts include forming partnerships with global CRO networks, AI research labs, and regulatory agencies. These collaborations will enable international deployment and

promote responsible AI use in clinical research. User feedback has been central to FiCRO's evolution. Pilot users highlighted the need for clearer cost transparency and better alignment with therapeutic areas. In response, FiCRO implemented a weighted scoring system and redesigned its user interface to improve matching accuracy and overall usability.

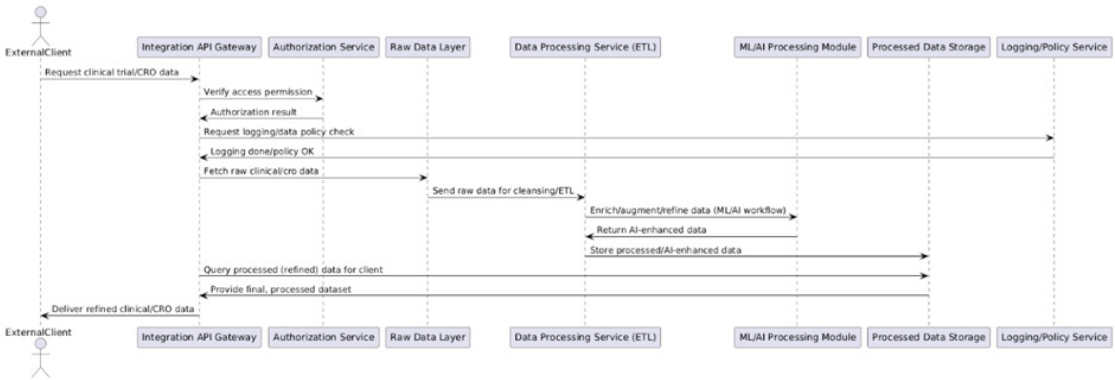
Additionally, learnings from medical device-specific documentation efforts (via the B company AX Project) informed the development of class-based protocol templates aligned with regional regulatory standards.

## 3 Use Case Requirements

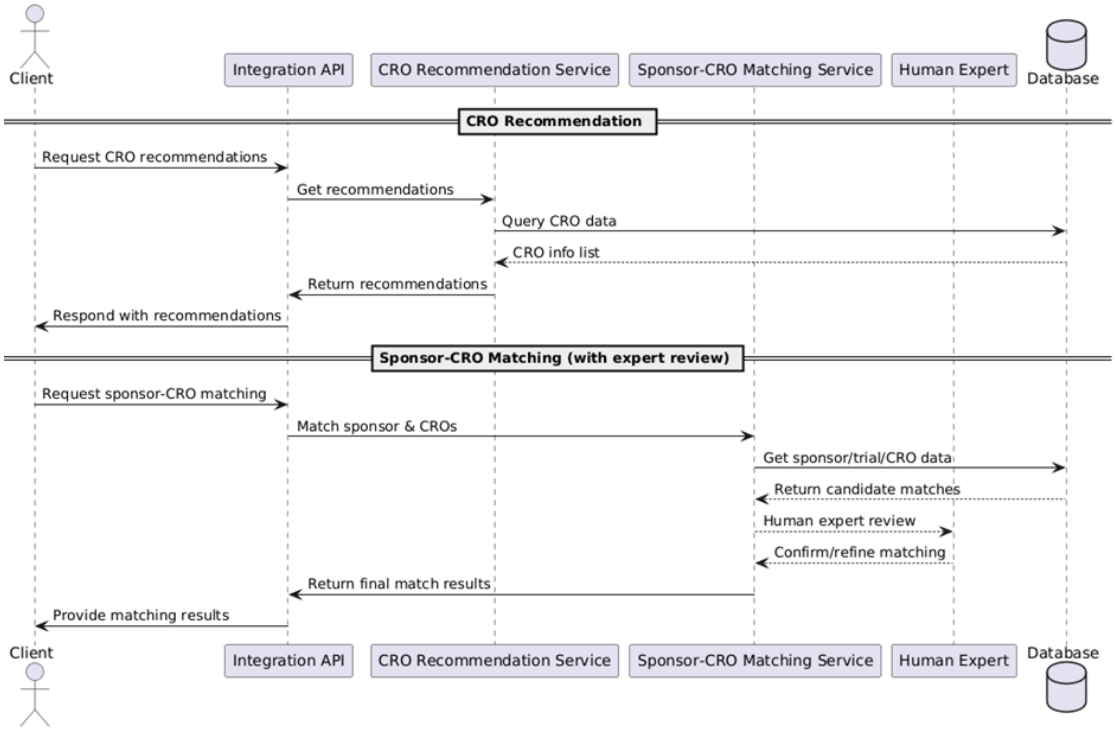
- **REQ-01:** The system must support access to anonymized and structured clinical trial data for AI model training and validation.
- **REQ-02:** It must include secure APIs for accessing and ranking CRO databases.
- **REQ-03:** The system must comply with international standards and regulations for data protection and clinical research, including GCP, GDPR and HIPAA. This ensures responsible data handling, ethical governance and auditability.
- **REQ-04:** The system must be deployed on a secure, cloud-based infrastructure capable of supporting AI model inference, storage and real-time processing across multiple geographic regions.
- **REQ-05:** The platform must provide an intuitive user interface and support seamless integration with existing Clinical Trial Management Systems (CTMS), Electronic Data Capture (EDC) systems and sponsor dashboards.
- **REQ-06:** The system must include a structured feedback mechanism to capture user inputs and real-world data, enabling iterative improvement of AI model accuracy, contextual relevance and explainability.

## 4 Sequence Diagram

### MediC (Main) – Functional Sequence Diagram



### FiCRO (Application/Service) Sequence Diagram




### Primary APIs Overview

| API Function Name | Description |
|---|---|
| getCRORecommendation | Provides recommendations of suitable CROs based on input clinical trial requirements and preferences. |
| matchSponsorWithCRO | Automatically matches sponsors and CROs using trial parameters and operational data, including expert review. |
| analyzeTrialFeasibility | Evaluates and predicts clinical trial feasibility, including site selection and patient recruitment analysis. |
| getResearchTrendInsights | Returns data-driven insights on the latest clinical research and development trends. |
| assessRiskCompliance | Assesses risk and compliance levels for CROs and trial sites using historical and regulatory data. |
| searchClinicalTrials | Allows searching the internal clinical trials dataset by various filters such as indications, phase, and region. |
| getCROProfile | Retrieves detailed information and analytics for a specific CRO. |

## 5 References

[1] ClinicalTrials.gov – U.S. National Library of Medicine. Referenced as a source of structured clinical trial protocol data.

[2] World Health Organization (WHO) ICTRP – International Clinical Trials Registry Platform. Consulted for data harmonization approaches and trial metadata fields.

[3] ICH-GCP E6(R2) – International Council for Harmonization. Reviewed for alignment with good clinical documentation practices, particularly in sponsor-CRO interactions.

[4] ITU-T Y.4208 – AI Capability Framework for Health Systems. Referred to as a conceptual framework for responsible AI integration in healthcare services.

[5] Citeline Trialtrove and Sitetrove (Norstella) – Commercial platforms referenced for benchmarking CRO capacity, performance, and trial feasibility analytics.

[6] Clarivate Cortellis – Used to cross-check market-level trends in clinical development and regional CRO distributions.

# Use Case 14: AI-Enabled Pathology Model for Precision Diagnosis and Treatment

**Organization:** Huawei Technologies Co., Ltd.

**Country**: China

**Contact Person(s)**:

Gao Hongli gaohongli@huawei.com

Wu Xiaomin amy.wuxiaomin@huawei.com

Liang Jiani liangjiani@huawei.com

## 1 Use Case Summary Table

| Item | Details |
|---|---|
| Category | Healthcare |
| Problem Addressed | According to the latest report released by the World Health Organization (WHO)'s International Agency for Research on Cancer (IARC), the global cancer burden is increasing, with 20 million new cancer cases and 9.7 million deaths worldwide by 2022 [1]. Early detection, early diagnosis and early treatment are the key. It is urgent to expand the accessibility of pathology diagnosis, improve the accuracy of pathology diagnosis and improve the level of pathology diagnosis at the grass-roots level.<br>In China, the pathology sector faces several challenges, including a significant shortage of qualified professionals—with fewer than 20,000 registered pathologists and an estimated gap of 70,000 to 140,000. Pathologists are unevenly distributed, with 70% concentrated in tertiary hospitals. Additionally, the compliance rate for initial diagnoses in primary hospitals is low, with only 13% showing complete compliance and around 30% achieving general diagnostic agreement. |
| Key Aspects of Solution | RuiPath model uses AI to analyze pathology slides, enabling faster and more accurate diagnosis across 11 subspecialties. The solution supports whole slide image analysis, clinical integration and reasoning-based diagnosis, reducing diagnosis time by 75% and improving early detection. From traditional microscopes to digitalization to intelligence, AI technology is used to achieve intelligent consultation and precise screening, improve the early diagnosis rate and treatment efficiency of common diseases, reduce misdiagnosis rates, help make up for the shortage of medical resources, align with international standards, and improve global public health standards. |
| Technology Keywords | Multimodal Large Language Model for Pathology; Self-supervised Learning; Clinical Decision Support Agent |

(continued)

| Item | Details |
|---|---|
| Data Availability | Private |
| Metadata (Type of Data) | Image, Electronic Medical Record (EMR) Report, Spatiotemporal Omics |
| Model Training and Fine-Tuning | • Mixture of Vision Encoder with Self-Supervised Learning (SSL)<br>• Vision-Language Alignment and Vision-Language Model Supervised Fine-Tuning (VLM SFT) |
| Testbeds or Pilot Deployments | Deployed internally at Ruijin Hospital [2], [3]. |
| Code repositories | Not yet available. A GitHub link will be added upon release. |

# 2 Use Case Description

## 2.1 Description

According to the latest report released by the World Health Organization's International Agency for Research on Cancer (IARC), the global cancer burden is increasing, with 20 million new cancer cases and 9.7 million deaths worldwide by 2022. Early detection, early diagnosis and early treatment are the key. It is urgent to expand the accessibility of pathology diagnosis, improve the accuracy of pathology diagnosis and improve the level of pathology diagnosis at the grass-roots level. For example, in China, the following problems exist in the pathology industry in the medical field: (1) There is a large gap in the number of doctors: no more than 20,000 registered pathologists, with a gap of 70,000-140,000. (2) Uneven distribution of pathologists: 70% are concentrated in third-class hospitals. (3) Low compliance rate of initial diagnosis: The complete compliance rate of initial diagnosis opinions in primary hospitals is only 13%, and the general compliance rate of initial diagnosis opinions is about 30%.

The DCS AI full-stack data storage solution has built the first multi-modality pathology model in China for automatic generation of clinical pathology reports. Pixel-level comparison, full slice traversal, interactive AI-assisted reading, and pathology report generation time were shortened from 40 minutes to 15 minutes, improving efficiency by 75%. From traditional microscopes to digitalization to intelligence, AI technology is used to achieve intelligent consultation and precise screening, improve the early diagnosis rate and treatment efficiency of common diseases, reduce misdiagnosis rates, help make up for the shortage of medical resources, align with international standards, and improve global public health standards.

In DCS AI full-stack data storage solution, the AI-Enabled Pathology Model named RuiPath model represents an innovative multimodal AI system designed for pathology analysis, integrating advanced machine learning techniques with large-scale medical data. Its technological approach was characterized by the following aspects:

Data Type, Structure, and Format:

- **Primary Data:** The core dataset consists of millions of digital pathology slide images (Whole Slide Images - WSIs). Initially, these images are largely unlabeled.
- **Supporting Data:** Tens of thousands of image-text captions, meticulously curated by medical experts (pathologists), linking specific visual features or regions within the pathology images to descriptive text. Thousands of image-based Chain-of-Thought (CoT) data instances. This structured data pairs pathology images with step-by-step reasoning text, designed to train the model in diagnostic thought processes.
- **Structure & Format:** The system leverages self-supervised learning algorithms on the vast corpus of pathology images to extract robust visual representations without requiring extensive initial manual annotation. The curated captions and CoT data provide structured, multimodal input crucial for aligning visual and textual features and enabling complex reasoning capabilities. Data processing, including WSI tiling/patching and anonymization, was managed by the ModelEngine which is a one-stop AI toolchain provided by DCS AI Solution, implying handling of standard WSI formats and subsequent generation of structured data formats suitable for model training (e.g., image patches linked to text via identifiers in formats like JSON or CSV).

I**mage Labelling System:** A semi-automated image labelling system was employed, facilitated by the Model Engine. This system assists in data engineering tasks, including efficient image annotation. While self-supervised learning reduced the dependency on exhaustive pixel-level labelling for initial feature extraction, expert-driven annotation was critical for creating the high-quality image-text captions and the image-based CoT data used for fine-tuning and aligning modalities.

**Knowledge Transfer and Update Format:** The system features a modular multimodal architecture comprising distinct components: a Visual Projector, an Image-Language Projector, and a Deep Reasoning Language Model module. Each module has well-defined input and output data format standards, allowing them to be trained independently (decoupled training). This modularity provides a clear format for knowledge transfer and updates. When new research findings, treatment mechanisms, or diagnostic criteria emerge, the knowledge base can be updated flexibly by two ways: 1) Incremental Training: Retraining specific modules with supplementary data incorporating the new knowledge (e.g., updating the language model with new medical texts, or the image-language projector with new image-caption pairs reflecting new findings). 2) Plugin Replacement: Replacing existing modules with newer versions incorporating updated algorithms or knowledge.

**Explanation of Prognosis/Diagnosis to Experts:** The integration of Chain-of-Thought (CoT) technology is key to providing explainability. The model explicitly generates and presents the reasoning steps leading to its diagnostic conclusions, mirroring a pathologist's thought process. This transparent reasoning pathway was presented to the expert pathologist. Furthermore, the system supports multi-turn, in-depth conversational interaction, allowing pathologists to query the model, ask follow-up questions about specific reasoning steps or image features, thereby significantly enhancing the interpretability and trustworthiness of the AI's suggestions. (Note: The source primarily mentions diagnosis and reasoning; explanation related to prognosis would follow similar principles if trained on relevant data).

**Feedback and Learnings from Ruijin Hospital Deployment:** The RuiPath model is currently deployed for pilot testing at Ruijin Hospital across 11 subspecialties, including breast, prostate, and thyroid pathology. Initial feedback based on performance metrics indicates promising results, with the system achieving over 90% accuracy in common tasks such as cancer subtype

classification and grading within these subspecialties. This deployment provides valuable real-world validation and data for further refinement.

**Use of Foundational Models:** The RuiPath model utilizes a foundational model. Specifically, DeepSeek R1 was employed, primarily through model distillation techniques during the training process. The deployment is based on DCS AI Solution, incorporating optimization techniques such as KV cache offloading and MoE (Mixture of Experts) expert prefetching. These optimizations have reportedly reduced the inference cost of the DeepSeek R1 component by 50%.

### The specific benefits of the solution are as follows:

1. **Breadth:** Covers 90% of the cancers in the world every year.
2. **Depth:** The depth of sub-speciality knowledge Q&A reaches the expert level, and the Q&A accuracy exceeds 90%, which is the leading level in the world.
3. **Efficiency:** In the past, it took about 10 minutes to diagnose a tissue slice under a microscope. Now, after AI interactive reading, AI calculation time for a single slice is only seconds, and the overall diagnosis efficiency is several times.

**Use Case Status:** The AI-Enabled Pathology Model is currently being pilot tested in 11 subspecialties at Ruijin Hospital, including breast, prostate and thyroid pathology. It will be officially implemented and operated at Ruijin Hospital in the latter half of 2025.

**Partners: Ruijin Hospital.**

## 2.2 Benefit of use case

The DCS AI solution provides the following benefits:

(1) Tool-based data engineering shortens the data preparation period for medical training by 80%.

(2) System-level model training and inference acceleration capabilities shorten the model training period by 30% and double the inference concurrency.

(3) Simplified application development platform, enabling non-professional developers to quickly get started. This programme embodies a commitment to industry, innovation and infrastructure, driving advances in cancer diagnostics and healthcare automation.

It leverages artificial intelligence technologies to deliver intelligent consultation and precise screening, improving the rate of early diagnosis and treatment efficiency of common diseases while reducing misdiagnosis rates. It helps cover the shortage of medical resources and raises global standards for public health.

In the realm of personnel efficiency, the transition from manual specimen information input and allocation of 2 staff members per lesion tissue to utilizing AI voice input and human verification has reduced the requirement to just 1 staff member per lesion tissue. Regarding diagnostic efficiency, initial reviewing physicians now spend 10-15 minutes per patient with AI-assisted image analysis and report editing, down from 20-40 minutes. Similarly, confirmatory physicians' time per patient has decreased to 10 minutes with AI prompting for review and automatic input, a significant improvement from the previous 15-30 minutes spent on manual review

and report modification. These innovations not only boost workforce productivity but also streamline diagnostic processes, aligning with the goals of promoting decent work practices and fostering economic growth.

DCS AI model project outlines a comprehensive model training process in two phases: first, compressing and encoding image data, and second, aligning image-text spaces. The model aims to enhance cancer diagnosis, facilitate similar medical history retrieval, automate pathology report generation, and drive innovation in education and research. By deploying this pan-cancer multi-modal pathology grand model, the goal is to achieve five core functionalities, leveraging cutting-edge technologies like Vision Transformers and multi-modal models. This initiative embodies a commitment to Industry, Innovation, and Infrastructure, driving advancements in cancer diagnostics and healthcare automation.

## 2.3 Future Work

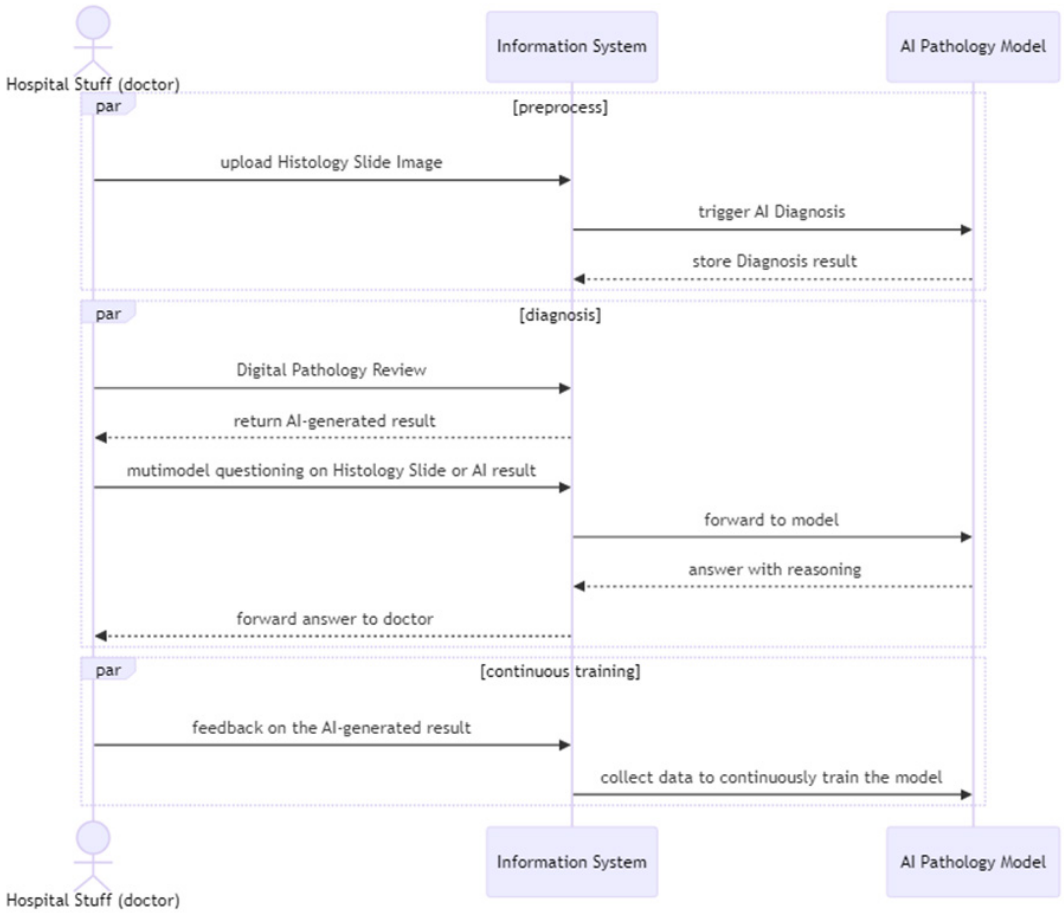
The development roadmap for the RuiPath system includes several key enhancements and expansions aimed at broadening its capabilities and impact:

- **Multimodal Data Integration:** Future iterations plan to move beyond the current reliance on H&E images and pathology report descriptions. The strategy involves incorporating additional data modalities, such as molecular diagnostics (e.g., genomic sequencing data) and immunohistochemistry (IHC) results. This integration aims to create a more comprehensive patient profile, enabling precision diagnostics and treatment planning across the patient's entire healthcare journey.
- **Expanded Downstream Task Coverage:** The scope of applications will be extended to include more complex downstream tasks beyond classification and grading. A primary focus will be the development and validation of robust prognosis prediction capabilities, leveraging the integrated multimodal data.
- **Second-Generation Model Architecture:** A next-generation architecture is under design to support these advancements. This includes:
    1. **Enhanced Self-Supervised Learning:** Updating the self-supervised algorithms to effectively process and learn from higher-resolution pathology images.
    2. **Advanced Multimodal Fusion:** Refining the Mixture of Experts (MoE) architecture to optimally fuse encoded information from diverse data modalities (H&E, molecular, IHC, etc.), aiming for further improvements in model accuracy and robustness.
- **Data Acquisition and Standardization:** Recognizing the critical role of data quality and diversity, future plans include:
    1. **Deepened Hospital Collaborations:** Establishing in-depth partnerships with multiple healthcare institutions.
    2. **Standard Development:** Collaborating with partner hospitals to jointly define standardized metrics and scanning protocols for digital pathology slides.
    3. **Multi-Center Data Collection:** Utilizing these standards to collect larger volumes of high-quality, multi-center data for model training, enhancing generalizability.
- **Open-Source Initiative and Expansion:** To foster broader adoption and innovation, there is a plan to open-source relevant aspects of the work in the second half of 2025. This initiative aims to empower more healthcare providers, particularly primary and community hospitals, with advanced AI tools for pathology analysis.

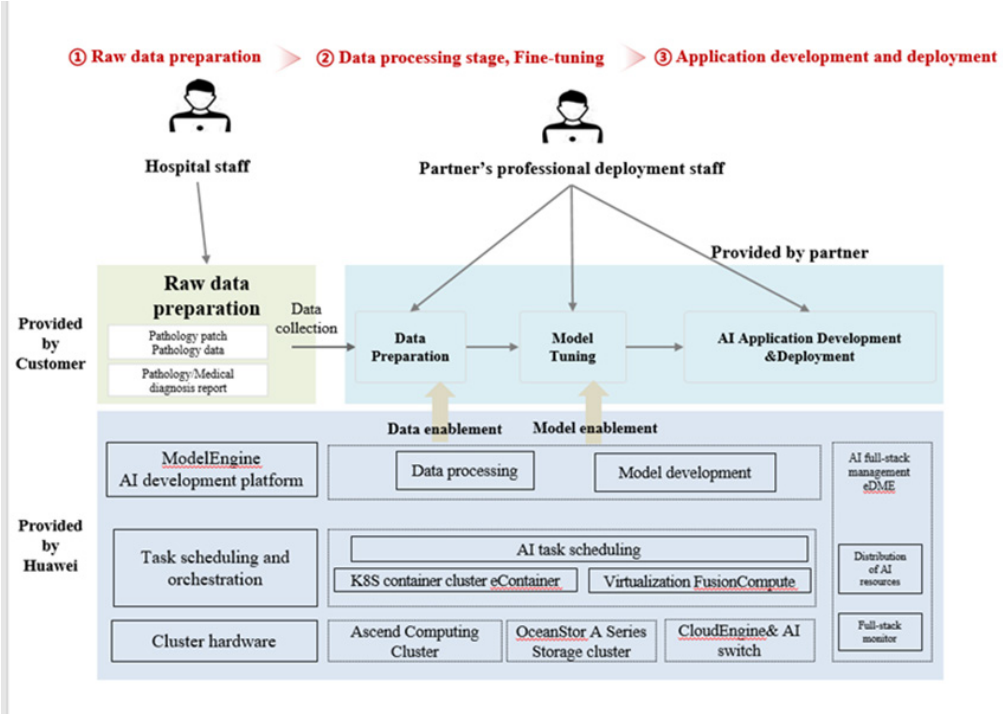
## 3 Use Case Requirements

- **REQ-01:** The system must connect to the hospital PIS (Pathology Information System) to automatically retrieve patients' WSI (Whole Slide Imaging), clinical history, and other relevant data, enabling comprehensive diagnostic assistance.
- **REQ-02:** The system must operate on a high-performance computing (HPC) cluster with low-latency interconnects and parallel file systems.
- **REQ-03:** The system must be updated when new research findings, treatment mechanisms, or diagnostic criteria emerge.

## 4 Sequence Diagram



Workflow Diagram




## 5 References

[1] "Global cancer burden growing, amidst mounting need for services", World Health Organization. [Online]. Available: https://www.who.int/news/item/01-02-2024-global-cancer-burden-growing--amidst-mounting-need-for-services [Accessed: 2025-05-27]

[2] "China's Ruijin Hospital, Huawei Launch AI Pathology Model", YICAI GLOBAL. [Online]. Available: https://www.yicaiglobal.com/news/chinas-ruijin-hospital-huawei-launch-ai-pathology-model [Accessed: 2025-05-27]

[3] "Digital Construction of Pathology Department of Ruijin Hospital: Reading 1000 Slide Images Per Second", HUAWEI. [Online]. Available: https://e.huawei.com/en/case-studies/storage/2023/ruijin-hospital-joint-oceanstor-pacific-digitalization-of-health [Accessed: 2025-05-27]

# Use Case 15: AI Specialist Diagnosis and Treatment Resources - Ophthalmology Model

**Organization**: Zhongshan Ophthalmic Center, Sun Yat-sen University

**Country**: China

**Contact Person(s)**:

Zhenzhe Lin, linzhenzhe@gzzoc.com

Mingjie Luo, zoc_mjluo@yeah.net

yuqiulin1@huawei.com

## 1 Use Case Summary Table

| Item | Details |
|---|---|
| Category | Healthcare |
| Problem Addressed | Pre-consultation, Decision Making, Medical Education, Medical Research, Hospital Management |
| Key Aspects of Solution | Mobile based solution, Multi-lingual Chatbot, Medical Reports, AI-Assisted Medical Image Diagnosis, Preliminary Diagnosis Eye Health Recommendations. |
| Technology Keywords | Retrieval-Augmented Large Language Model (RAG-LLM), Machine Learning (ML), Visionome, API |
| Data Availability | Private |
| Metadata (Type of Data) | Text, Audio, Visual |
| Model Training and Fine-Tuning | Pretrain with Medical data, SFT |
| Testbeds or Pilot Deployments | Pilot initiatives were implemented across multiple regions in China. |
| Code repositories | Not yet available. |

# 2 Use Case Description

## 2.1 Description

ChatZOC Ophthalmology Large Model is a specialized medical AI model designed by Zhongshan Ophthalmic Center (ZOC), Sun Yat-sen University. It is to enhance the intelligence and efficiency of ophthalmic diagnosis, patient services, scientific research, medical education, and hospital management. By leveraging Natural Language Processing (NLP) and deep learning technologies in image analysis, ChatZOC offers intelligent consultations, precise screening, diagnostic assistance, and research support.

ChatZOC pioneered dense anatomical annotation for ophthalmic medical images, significantly improving data utilization. A self-developed Visionome model can recognize 22 anterior segment pathologies and provide diagnostic recommendations by integrating visual and textual data. On the language side, ChatZOC employs multi-level knowledge retrieval and patient intent recognition to deeply understand complex clinical questions and perform medical reasoning. It was also implemented an intelligent question rewriting mechanism to enhance input precision and improve response interpretability.

It effectively addresses critical challenges such as the shortage of ophthalmologists and healthcare disparities between urban and rural areas. Compared to traditional, labor-intensive diagnostic approaches, ChatZOC enables rapid responses, standardized diagnostic protocols, and large-scale screening capabilities. However, there remains room for improvement in recognizing complex conditions, providing personalized treatments plans, and managing high computational demands. Moving forward, ChatZOC will continue to evolve and innovate, contributing to the sustainable development of global eye health.

**Supplemental Information:**

In response to the questions raised by the judges of ITU, we provide the following information

1) how to include expert supervision and audit?

Response: Firstly, regarding the data aspect, we utilized professional data that was rigorously standardized by professional doctors. In terms of the model, we established a RAG (Retrieval Augmented Generation) database and have corresponding supervision mechanisms for the output. Expert supervision is integral to the ChatZOC framework. ChatZOC serves as an assistive tool designed to support clinicians, not replace their judgment. Ophthalmologists retain final decision-making authority and perform audits on AI-generated diagnostic reports to ensure clinical safety, accuracy, and appropriate application.

2) how is the preparation and labelling of data done in this framework?

Response: The data used for training is medical data from real doctor-patient interaction scenario. Before commencing annotation, a clear and specific annotation guideline needs to be formulated to ensure a common understanding of the standards among all personnel involved. To minimize the influence of personal bias, a double-blind annotation method can be employed, where each data point is independently annotated by two junior doctors, followed by review and arbitration by a senior doctor with 20 or more years of experience, ensuring annotation consistency and accuracy. Upon completion of annotation, the data also undergoes validation and correction to ensure its completeness and accuracy.

3) what type of foundation model is used? e.g. DeepSeek or GPT, etc

Response: The language model component of ChatZOC was independently developed by the ZOC algorithm team. While it currently utilizes the Qwen 2.5 open-source architecture as its foundation, the overall system is designed for flexibility and compatibility with multiple large language model bases. This architecture ensures ChatZOC is not solely dependent on a single foundation like Qwen 2.5 and allows for potential migration to other bases as needed. Crucially, regardless of the base model, the system is heavily fine-tuned using ZOC's proprietary medical knowledge and extensive ophthalmic case data, creating a highly specialized tool tailored for eye care.

4) are there references or articles or blogs or academic papers that can be referenced?

Response: The key underlying techniques like the Retrieval-Augmented Large Language Model (RAG-LLM)[2] and Visionome[3] have been documented in publications such as JAMA Ophthalmology and Nature Biomedical Engineering respectively.

5) are there open datasets or models?

Response: At present, the specific datasets compiled and the ChatZOC models developed are not publicly available. This is due to strict patient data privacy laws and prevailing regulatory constraints governing medical data.

6) could you add feedback or learnings from the deployments or pilots?

Response: Through previous pilot deployments, we have also identified areas for further model optimization, such as designing it to be multilingual to meet the needs of users speaking different regional dialects; the system needs to understand patients' anxiety related to their illness and enhance its capacity for humanistic care; and for users without a medical background, the language used in responses needs to be easily understandable.

Use Case Status: The use case is part of a larger research project.

Partners: Huawei Technologies Co., Ltd.

## 2.2 Benefit of use case

ChatZOC leverages artificial intelligence technologies to deliver intelligent consultations, precise screening, and diagnostic assistance, significantly improving the rate of early diagnosis and treatment efficiency of ophthalmic diseases while reducing misdiagnosis rates. It helps cover the shortage of medical resources and raises global standards for eye health. Additionally, with the support of telemedicine technologies, patients can also access high-quality ophthalmic medical services remotely, ensuring better health outcomes for more individuals.

AI large models can provide medical knowledge Q&A, virtual case-based teaching, and clinical decision support, helping medical students and young doctors efficiently acquire ophthalmic diagnostic and treatment skills, thus advancing the digital and intelligent development of medical education.

Integrating cutting-edge artificial intelligence technologies with medicine, promote the construction of smart medical infrastructure, and optimizing ophthalmic service processes,

thereby providing technical support and a demonstrative effect for technological innovation and sustainable development in the healthcare industry.

ChatZOC can remotely provide high-quality ophthalmic medical services, overcoming geographical and resource limitations, allowing patients in remote and resource-poor areas to access quality healthcare. This helps narrow the healthcare service gap between urban and rural areas and reduces healthcare inequalities.

## 2.3 Future Work

### a. Data Collection

Continuously expand and optimize ophthalmic medical datasets to encompass clinical data from diverse regions, age groups, and various eye diseases, ensuring data comprehensiveness and representativeness. Strengthen collaboration with international ophthalmic medical institutions to acquire diverse global data, enhancing the model' s applicability and accuracy among different populations and scenarios.

### b. Proof of Concept Development

In the future, clinical scenarios will be validated, to simulate real clinical consultation, screening and treatment scenarios. Multimodal diagnostic integration will be conducted by combining ophthalmic imaging data, such as fundus images and OCT scans.

### c. Model Development

Based on the existing ChatZOC ophthalmology large model, We will optimize the ChatZOC ophthalmology model to enhance its reasoning and clinical decision support, focusing on complex cases and rare diseases. A lightweight version will also be developed for deployment in resource-limited areas and on mobile devices.

### d. Create New Variations/Extensions to the Same Use Case

Building on core functions such as intelligent consultations, precise screening, and diagnostic assistance, we aim to explore additional application scenarios, including remote surgical guidance, personalized treatment recommendations, and patient health management. Furthermore, we plan to expand into other subspecialties, such as pediatric ophthalmology and geriatric ophthalmology, to meet the specific demands of diverse patient groups.

### e. Standards Development Related to the Use Case

In the future, efforts will be made to establish industry standards and technical guidelines for AI-assisted ophthalmic diagnosis and treatment, ensuring that ChatZOC's diagnostic recommendations, data utilization, and result presentation align with international standards. Legal and regulatory requirements regarding patient data privacy and model usage security will be clearly defined, along with the establishment of a robust safety review mechanism for AI models to ensure safety and reliability in clinical use.

**f. Setup Reference tools, Notebooks and Simulation Environment**

A standardized model simulation and testing environment will be established, along with the creation of an ophthalmic AI diagnostic simulation platform to test and optimize real-world medical scenarios. Provide detailed technical documentation and operation instructions to lower the technical entry barrier.

## 3 Use Case Requirements

REQ-01 (Functional): The system must accurately analyze diverse ophthalmic imaging data (e.g., fundus photographs, OCT scans) to detect key pathological signs associated with common preventable or treatable blinding eye diseases such as Diabetic Retinopathy (DR), Glaucoma, Age-related Macular Degeneration (AMD), and Pathological Myopia.

REQ-02 (Functional): The system must provide clinicians with clear, evidence-based outputs, including diagnostic likelihoods, disease severity grading (where applicable), identification of critical findings requiring urgent referral, and potentially risk stratification for disease progression.

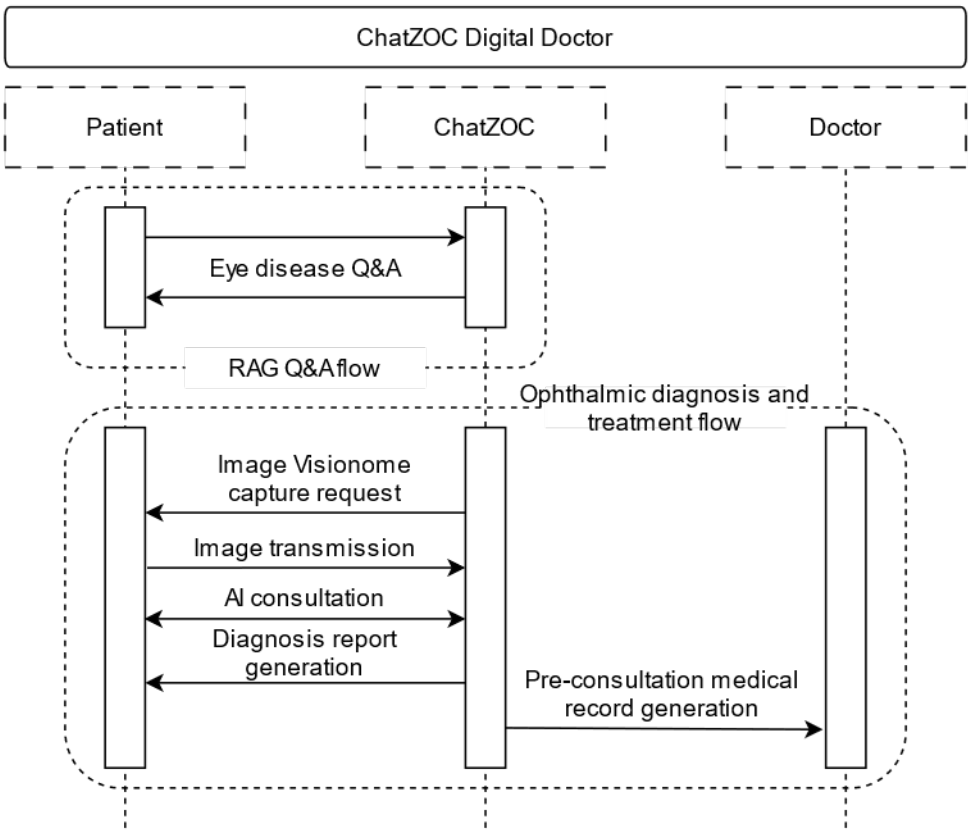
REQ-03 (Functional): The system must support multi-modal data input, capable of processing not only images but also relevant clinical text information (e.g., patient history snippets, symptoms described) to provide a more contextualized assessment.

REQ-04 (Functional): The system must offer a natural language interface allowing clinicians to interactively query for differential diagnoses, ask specific questions about findings, or retrieve relevant information pertaining to the case, mimicking consultation with an expert.

REQ-05 (Operational): The system must ensure robust data privacy and security, complying with relevant healthcare data protection regulations through measures like data anonymization, encryption, and secure access controls.

REQ-06 (Technical): The underlying AI models must be rigorously validated on diverse, representative datasets for accuracy, fairness, and robustness across different populations and equipment types.

## 4 Sequence Diagram




- Insert a sequence diagram or conceptual diagram clearly illustrating interactions between different components or users within the use case.

## 5 References

[1] ChatZOC. “中山大学中山眼科中心联合华为打造国产大模型ChatzOC，用AI提高诊疗效率.” Accessed June 24, 2025. https://chatzoc.gzzoc.org.cn.

[2] Zhang, Y., X. Liu, W. Li, et al. “Predictive Value of Artificial Intelligence in Diagnosing Pancreatic Cancer Using CT Imaging: A Multicenter Study.” PubMed, published January 2024. https://pubmed.ncbi.nlm.nih.gov/39023885/.

[3] Wang, H., J. Chen, M. Zhou, et al. “Integration of Large Language Models in Clinical Decision Support Systems: A Retrospective Evaluation.” PubMed, published February 2024

[4] Huawei. “中山大学中山眼科中心与华为合作推出ChatzOC大模型：AI辅助诊疗新时代.” Huawei Newsroom, June 2024. https://www.huawei.com/cn/news/2024/6/chatzoc-ai.

[5] Luo M, Pang J, Bi S, et al. Development and Evaluation of a Retrieval-Augmented Large Language Model Framework for Ophthalmology. JAMA Ophthalmol. 2024;142(9):798–805. doi:10.1001/jamaophthalmol.2024.2513

# Use Case 16: Traditional Chinese Medicine Industry Leigong Large Model

**Organization**: Huawei Technologies Co., Ltd. & Yunnan Baiyao Group Co., Ltd.

**Country**: China

**Contact Person(s)**:

Zhang Yi, lisaa.zhang@huawei.com

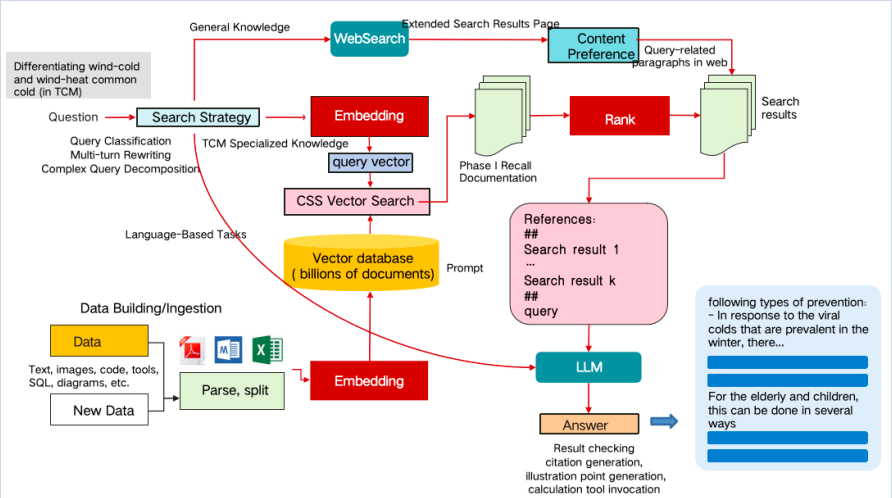
Meng Deli, mengdeli@ynby.cn

Dang Yabin, dangyabin@ynby.cn

## 1 Use Case Summary Table

| Item | Details |
|---|---|
| Category | Healthcare |
| Problem Addressed | 1. Enhancing data quality in TCM R&D processes.<br>2. Mitigating forgetting in Large Model incremental training.<br>3. Developing AI applications for intelligent identification, sorting, cultivation guidance, real-time growth monitoring, and public health education to elevate quality control and standardization levels of TCM. |

(continued)

| Item | Details |
|---|---|
| Key Aspects of Solution | 1. Using embedding technology, we developed a system for parsing, storing, and retrieving multimodal data. By converting TCM prescriptions, literature, and related materials into vector representations, we built a multimodal (e.g., text, images, PDFs) vector knowledge base, effectively addressing challenges in data utilization.<br>2. By incorporating KooSearch-enhanced retrieval, we have developed core capabilities such as query rewriting, post-retrieval risk validation, citation localization, and multimodal (text-image) match positioning—effectively reducing hallucinations in large language model outputs.<br>3. Integrating large models and small models, we developed a multimodal RAG platform for fast, scalable AI assistant deployment across diverse use cases.<br>  |
| Technology Keywords | Digital foundation, knowledge graphs, Vector database, Large Language Model |
| Data Availability | TCM prescriptions, TCM literature, Classical TCM Texts, Chinese Pharmacopoeia, Phytopharmaceutical Study Datasets, Labeled TCM Practice Data |
| Metadata (Type of Data) | text, image |
| Model Training and Fine-Tuning | Huawei Pangu's Large Language Model solution provides an AI computing infrastructure, AI development platforms, and tools to construct a vector knowledge database. Yunnan Baiyao collects high-quality datasets, trains and fine-tunes them on the Pangu Large Model to develop the Lei Gong Large Model. By integrating prompt engineering, it provides comprehensive application services: supporting farmer cultivation, accelerating product R&D, and delivering health management solutions. |
| Testbeds or Pilot Deployments | Leigong Large Model for Traditional Chinese Medicine (TCM)[5] |
| Code repositories | Not available |

# 2 Use Case Description

## 2.1 Description

- **Objectives**: Yunnan Baiyao is committed to optimizing the entire industry chain through AI technology, achieving digitalization, visualization, and traceability throughout the entire value chain, from medicinal herb cultivation to R&D, production, and service. Utilizing the multi-dimensional capabilities of large models, such as knowledge retrieval, agile computing, and interactive experiences, it is applied in various scenarios including traditional Chinese medicine (TCM) R&D, decision support, knowledge popularization, and digital marketing. It provides services for diverse groups: It enables health-conscious individuals to access wellness guidance, patients to obtain disease management and basic medication instructions, Chinese herbal medicine growers to enhance cultivation techniques, sales professionals to master product knowledge, primary TCM practitioners to receive diagnostic support, and students to acquire specialized knowledge in TCM.
- **Pain point**: The quality of TCM R&D data is uneven. The large models are prone to forgetting during incremental training. Data is scattered and insufficiently used. Traditional methods make it difficult to achieve in-depth data mining of the whole industry chain. Lack of ability to handle complex knowledge and scenarios.
- **Solution**: Based on the high-quality data accumulated by Yunnan Baiyao, Leigong large model learns tens of millions of high-quality TCM documents, millions of structured prescriptions and TCM drug knowledge graphs, and builds a vector knowledge base. With Huawei Pangu's AI computing power foundation, AI platform, and development tools, it has brought about the sharing and integration of AI technologies around the world.
- The Chinese herbal medicine industry has deployed AI applications, including intelligent identification and sorting—achieving 30-50x efficiency gains and 300% accuracy improvements over traditional methods. AI-driven planting guidance and smart cultivation monitoring have reduced annual return costs for Poria cocos ( ) alone by over 50 million CNY. These advancements elevate quality control and standardization, strengthening global trust and acceptance of traditional Chinese medicines.
- Chinese medicine assists decision-making, completes model training (such as incremental pre-training), data processing (ancient, semi-white and semi-ancient), data proportioning, and data synthesis applications, and solves problems such as data quality and large model incremental training forgetting in the research and development process of Chinese medicine.
- The knowledge science service enables comprehensive querying and interpretation of TCM knowledge (including composition of medicinal materials and symptom explanations), improving knowledge acquisition efficiency by 3-5 times. With its traceable knowledge system and rigorous verification mechanism, the service significantly reduces hallucination issues in large language models, increases effective knowledge utilization rate, and enhances patient satisfaction by 30-50%. This approach effectively addresses international misunderstandings and biases towards TCM, facilitating the global dissemination and cultural exchange of traditional Chinese medicine.

**Use Case Status:**

The use case is part of a larger research project. Currently, the intelligent TCM herb recognition system, automated herb sorting system, and TCM knowledge science application have been developed and deployed. We regularly analyse problem feedback for project iteration and optimization and collaborate with institutions such as the Chinese Academy of Medical Sciences and the University of Traditional Chinese Medicine to continuously improve and revise the training data for large models.

Partners:

**Yunnan Baiyao Group:** Focus on various scenarios including traditional Chinese medicine (TCM) R&D, decision support, knowledge popularization, and digital marketing.

**Huawei:** Provides unified digital infrastructure, and Pangu Large Language Model solution.

## 2.2 Benefit of use case

- Addressing patients' conditions, leveraging learning outcomes from over 10 million pieces of TCM literature and millions of structured prescriptions, enhance the judgment capabilities in complex scenarios such as TCM disease reasoning, providing precise drug recommendations and treatment plans. Answer public questions about TCM, improve public health literacy, and help people better prevent diseases and manage their health.
- Through the multi-modal capabilities of the Leigong large model, quickly identify the growth status of Chinese medicinal materials, prevent the impact of pests and diseases, and provide professional solutions; farmers can follow the intelligent planting GAP (Good Agricultural Practice) guidance at any time. By solving challenges such as data quality and incremental training forgetting in TCM R&D, accelerate the development process of new drugs and therapies, and promote the innovative development of TCM.

## 2.3 Future Work

The future work of TCM Industry Leigong Large Model includes:

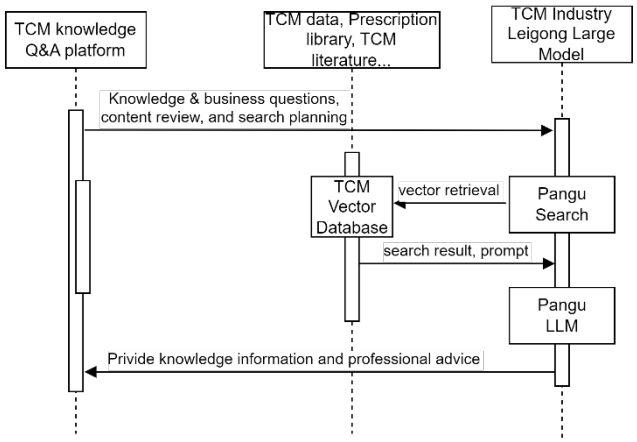
1. Innovative application scenarios: Improve model development in three types of scenarios: intelligent planting, intelligent diagnosis, and intelligent supply chain, such as intelligent planting scenario tools, intelligent diagnosis of Chinese medicinal materials growth, and intelligent consultation for TCM.
2. High-quality dataset: Build high-quality training and evaluation datasets for the scenario-based solutions of TCM industry and medical health services. Regularly update the datasets to adapt to the evolving intelligent technology and market demand changes in TCM.
3. Continuous operation: Establish a trusted data space for sustainable operations, operational organization structure, and resource allocation mechanism.
4. Industry standards: Develop standards covering the application specifications, data models, data dictionaries, data security, and privacy protection guidelines for AI models in the cultivation and quality of TCM, medical health services, and health knowledge popularization, laying a foundation for the healthy and orderly development of the industry. Data models and data dictionaries standards will be developed in ITU-T SG20, while TCM industry AI model application specifications will be developed in ITU-T SG21.

# 3 Use Case Requirements

- **REQ-01:** Constructing a Traditional Chinese Medicine (TCM) vector database based on TCM literature, prescriptions, and other TCM-related data. The TCM vector database is being developed in a phased and compliant manner. Currently, it is in a restricted access mode, open only to partner customers. It will be fully opened in the future.
- **REQ-02:** Yunnan Baiyao leverages its compiled high-quality datasets to train and fine-tune the Pangu Large Model, developing the Lei Gong Large Model specialized for the TCM industry.

- **REQ-03:** By integrating the TCM vector database with prompt engineering, the system mitigates hallucination during incremental training and inference processes of large models.
- **REQ-04:** Expert review and annotation: During the knowledge database construction phase, only authoritative source content is selected, such as official authoritative texts like the "Chinese Pharmacopoeia" and "Chinese Herbal Medicine." A team of TCM experts annotates key information such as the efficacy of medicinal materials and contraindications, and establishes a "confidence level grading" system.
- **REQ-05:** Quality control of model training and output: A dual review mechanism is used to evaluate the output results. First, machine filtering of keywords is employed, such as automatically blocking high-risk responses with absolute statements like "cure" and "anti-cancer". Second, manual spot checks are conducted, with 5% of the generated results randomly selected daily for evaluation by professional Chinese medicine practitioners.
- **REQ-06:** Adhering to China's "Good Agricultural Practices for Chinese Medicinal Materials Certification Management Measures" and "Interim Measures for the Management of Generative Artificial Intelligence Services." The Leigong large model complies with requirements for data security, content accuracy, and qualification filing. In data security, all user health data is anonymized; for content accuracy, content is reviewed by a team of traditional Chinese medicine experts, RAG technology is used to reduce LLM hallucinations, output results are marked with content sources, and labeled as "Content is for reference only and does not replace professional diagnosis and treatment." For qualification filing, it is stated that the current version is a tool to assist doctors, not a replacement for diagnosis, and does not require medical qualification filing.

## 4 Sequence Diagram




Based on data from traditional Chinese medicine, prescription libraries, and traditional Chinese medicine literature, transform it into vector data to construct traditional Chinese medicine vector data.

Initiate knowledge and business inquiries on the platform, first review the content and plan the search. Then use the Pangu search function to achieve natural language vector retrieval. For search results, combine prompt engineering to automatically summarize, providing knowledge information and professional advice. Accelerate business processing flows, reduce large model

hallucination issues, effectively enhance problem understanding, and improve the accuracy of answers.

- Cyberspace **Administration of China.** "Notice by the Cyberspace Administration of China on Issuing the Interim Measures for the Management of Generative Artificial Intelligence Services." Published July 13, 2023. https://www.cac.gov.cn/2023-07/13/c_1690898327029107.htm.

## 5 References

[1] Cyberspace Administration of China. "Notice by the Cyberspace Administration of China on Issuing the Interim Measures for the Management of Generative Artificial Intelligence Services." Published July 13, 2023. https://www.cac.gov.cn/2023-07/13/c_1690898327029107.htm.

[2] State Council of the People's Republic of China. "Decision of the Central Committee of the Communist Party of China and the State Council on Further Strengthening Rural Health Work." Published 2004. https://www.gov.cn/gongbao/content/2004/content_63083.htm.

[3] Yunnan Baiyao. "Yunnan Baiyao Group Unveils the 'Leigong Large Model.'" Accessed June 18, 2025. https://www.yunnanbaiyao.com.cn/view/ynbyPc/1/92/view/7804.html.

[4] Huawei Enterprise. "Group Cloud Empowers Yunnan Baiyao's Intelligent Manufacturing." Accessed June 18, 2025. https://e.huawei.com/cn/case-studies/industries/manufacturing/202501-group-cloud-yunnanbaiyao.

[5] Welkin AI. "Yunnan Baiyao Leigong Large Model Platform." Accessed June 18, 2025. https://ynbytcm.welkinai.com.

# Use Case 17: MoralQuest: AI-Powered Gamification of Indian Moral Tales for Neurodiverse Social Development in children with Autism Spectrum Disorder (ASD)

**Organization**: Christ University, Bangalore, India

**Country**: India

**Contact Person(s)**:

Gobi Ramasamy, gobi.r@christuniversity.in

Balakrishnan C, balakrishnan.c@christuniversity.in

Priyansha Upadhyay, priyansha.upadhyay1304@gmail.com

Satyam Kumar, satyamkumar0905@outlook.com

## 1 Use Case Summary Table

| Item | Details |
|---|---|
| Category | Healthcare |
| Problem Addressed | Current interventions for children with Autism Spectrum Disorder (ASD) often fall short because they're static and generic. They don't adapt to each child's individual learning pace and evolving needs, leading to limited effectiveness in fostering social development. |
| Key Aspects of Solution | AI-powered gamified mobile application using traditional Indian moral tales, Personalized interactive quests adapting to learning pace and emotional state, Natural Language Processing (NLP) for content adaptation, Predictive analytics to forecast challenges and maintain engagement, Dynamic difficulty adjustment, Behavioural data collection and analysis, Flexible reward mechanisms (points, badges, leaderboards), Caregiver insights and feedback system |
| Technology Keywords | NLP, Predictive Analytics, Gamification, Adaptive Learning, Behavioural Analysis |
| Data Availability | Private (in-app interaction data, periodic questionnaires) |
| Metadata (Type of Data) | Textual (behavioural, Cognitive, User interaction), Visual, Auditory, Performance metrics |

(continued)

| Item | Details |
|---|---|
| Model Training and Fine-Tuning | 1) Machine learning models for personalization and adaptivity:<br>• Classification models (Random Forests) to categorize interaction patterns or infer emotional states from interaction proxies.<br>• Regression models to predict engagement scores or learning trajectory.<br>• Contextual bandit algorithms for optimizing the sequence of quests and content presentation.<br>2) NLP for content processing and generation:<br>• Sentiment analysis on caregiver feedback to identify areas of concern or success.<br>• Text simplification algorithms to adjust readability of instructions and story elements.<br>3) Predictive analytics for engagement forecasting:<br>• Sequence modeling (ex: LSTMs, Transformers if data volume allows, otherwise simpler Markov models) on interaction logs to predict when a child might become disengaged or frustrated.<br>4) Reinforcement learning for reward optimization:<br>• Models (Q-learning or policy gradient methods like PPO, with appropriate simplifications for the domain) will be trained to optimize the delivery (timing, type, magnitude) of rewards. The reward function for the RL agent will be a composite function aiming to maximize sustained engagement and learning progress, potentially including terms for:<br>• + (Quest Completion Success)<br>• + (Correct Application of Moral in Scenario)<br>• – (Hints Used Penalty)<br>• + (Normalized Engagement Time within Optimal Range)<br>• – (Signs of Frustration/Disengagement from interaction patterns) |
| Testbeds or Pilot Deployments | In development, the Minimum Viable Product (MVP) phase is currently underway. The MVP will have the following features.<br>1. Basic gamified mobile interface with simple touch interactions<br>2. A limited library of traditional Indian moral tales converted into interactive quests<br>3. Fundamental AI personalization using basic classification models for behavioral patterns<br>4. Simple dynamic difficulty adjustment based on rule-based thresholds (ex: X correct answers increases difficulty)<br>5. Basic reward system (points, badges) without advanced AI optimization<br>6. Essential data collection from user interactions<br>7. Basic caregiver dashboard with simple progress metrics<br>8. Multi-modal content delivery (visual illustrations, audio narration, text) |
| Code repositories | Not publicly disclosed |

# 2 Use Case Description

## 2.1 Description

MoralQuest introduces a transformative AI-powered gamified mobile application that leverages traditional Indian moral tales to enhance social behaviour in children with Autism Spectrum Disorder (ASD) [1]. The solution integrates advanced machine learning and Natural Language Processing (NLP) to deliver personalised interactive quests that adapt to each child's learning pace and emotional state. Predictive analytics forecasts potential challenges to maintain optimal engagement and foster targeted learning outcomes, dynamically adjusting difficulty levels. To ensure accessibility and engagement for children with diverse sensory profiles, MoralQuest employs a multi-modal delivery approach. Content is presented through a combination of engaging visuals (illustrations, simple animations), clear auditory components (voice narration, sound effects, optional music), and supportive text (on-screen instructions, optional subtitles). Key interactions leverage intuitive touch-based controls. This combination allows the AI to potentially personalize not just the content difficulty but also the preferred mode of presentation for each child. Existing interventions typically rely on static, one-size-fits-all content that fails to evolve with a child's unique needs & growth rate, resulting in disengagement and limited progress. Traditional methods lack real-life feedback, personalization, and ongoing behaviour assessment, which are key factors in the social and cognitive development of children with ASD [2].

This AI-powered approach continuously collects and analyses behavioural data from in-app interactions as well as through periodic questionnaires. This data enables precise tailoring of challenges and rewards, with a flexible mechanism integrating points, badges, and leaderboards to prompt further development. The system refines cognitive, emotional, and social skills and provides caregivers with actionable insights to support each child's development. The core learning mechanism involves translating the morals embedded within these narratives into concrete, actionable social skills. For example, a story about honesty might generate interactive quests focusing on understanding truthfulness versus lying, recognizing associated emotions (guilt, trust), and predicting social consequences within simplified scenarios derived from the tale. The AI personalizes these quests to target specific social development goals relevant to ASD, such as perspective taking, emotion recognition, or understanding social cues, using the familiar and engaging context of the story [2].

**Use Case Status:** The use case is part of a larger product development

## 2.2 Benefit of use case

This AI-enhanced learning platform is specifically designed for children with Autism Spectrum Disorder. Rather than offering a one-size-fits-all approach, the solution adapts to each child's unique learning journey, providing personalized challenges and constructive feedback at every step. By making use of traditional Indian moral stories in the educational experience, the platform helps children build academic skills and a strong sense of cultural identity and social values.

The use case provides timely emotional support and interventions tailored to each child's needs and makes sophisticated educational technology accessible to children from all backgrounds.

The initiative provides an educational tool and a stepping stone to more significant equity and opportunity for children with autism by thoughtfully utilizing the transformative power of artificial intelligence with culturally rich content.

## 2.3 Future Work

Our path forward will focus on developing a much stronger, expandable framework to enhance the technological capabilities and worldwide reach of this AI-powered learning platform for children with Autism Spectrum Disorder. The strategic direction encompasses several key areas:

- More Robust Data Collection

  A sophisticated system will gather behavioural, cognitive, and emotional information from children's interactions with the platform, feedback from parents and teachers, and potentially from wearable devices. This approach will be implemented with strict attention to privacy regulations like India's Digital Personal Data Protection Act (DPDP) 2023 amongst others, for local and GDPR abroad, creating a valuable foundation for refining the AI models and personalising each child's learning experience.

- Extensive Real-World Testing

  In partnership with autism centres, schools, and clinical specialists, an extensive pilot program will test the platform in everyday settings. This critical phase will allow continuous improvements based on feedback from diverse users and performance measurements. The insights gained will help validate the approach's effectiveness and scalability.

- Advancements in AI Capabilities

  The next product development stage will incorporate cutting-edge machine learning techniques, including reinforcement learning to create more responsive reward systems, neural networks to anticipate behavioral patterns, and natural language processing to analyze engagement levels. Regular updates based on incoming data will ensure the platform remains responsive and personally relevant to each child.

- Expanding Accessibility

  The team will include multiple regional languages and immersive augmented or virtual reality experiences to broaden the impact. These additions will make the platform more culturally adaptable and enhance its appeal and usability across different global contexts.

- Setting Industry Standards

  A significant goal is to establish benchmarks and best practices for AI-driven special education. This includes developing open-source tools, detailed implementation guides, and testing environments to encourage collaborative research and smooth integration with existing educational systems.

# 3 Use Case Requirements

- **REQ-01**: It is critical that the AI-powered content adaptation system is capable of personalising story-based quests
- **REQ-02**: It is critical that machine learning algorithms (classification, regression, sequence modelling) be used for analysing user behaviour and learning patterns.
- **REQ-03**: It is critical for natural language processing capabilities to adapt to content complexity
- **REQ-04**: It is critical that a multi-modal content delivery supports visual (images, animations), auditory (voice, sounds), and textual elements, adaptable to user needs.

- **REQ-05**: It is critical that a traditional Indian moral tales content library with educational significance is added.
- **REQ-06:** It is critical that a Dynamic difficulty adjustment mechanism responsive to user performance is integrated
- **REQ-07**: It is expected that a Predictive analytics system to forecast engagement challenges is incorporated
- **REQ-08**: It is expected that a comprehensive data collection from in-app interactions and questionnaires
- **REQ-09**: It is expected that flexible gamification elements, including points, badges, and leaderboards, are incorporated.
- **REQ-10**: It is expected that a user-friendly mobile interface appropriate for children with ASD
- **REQ-11**: It is of added advantage that the dashboard for caregivers provides actionable insights on child development

## 4 Sequence Diagram

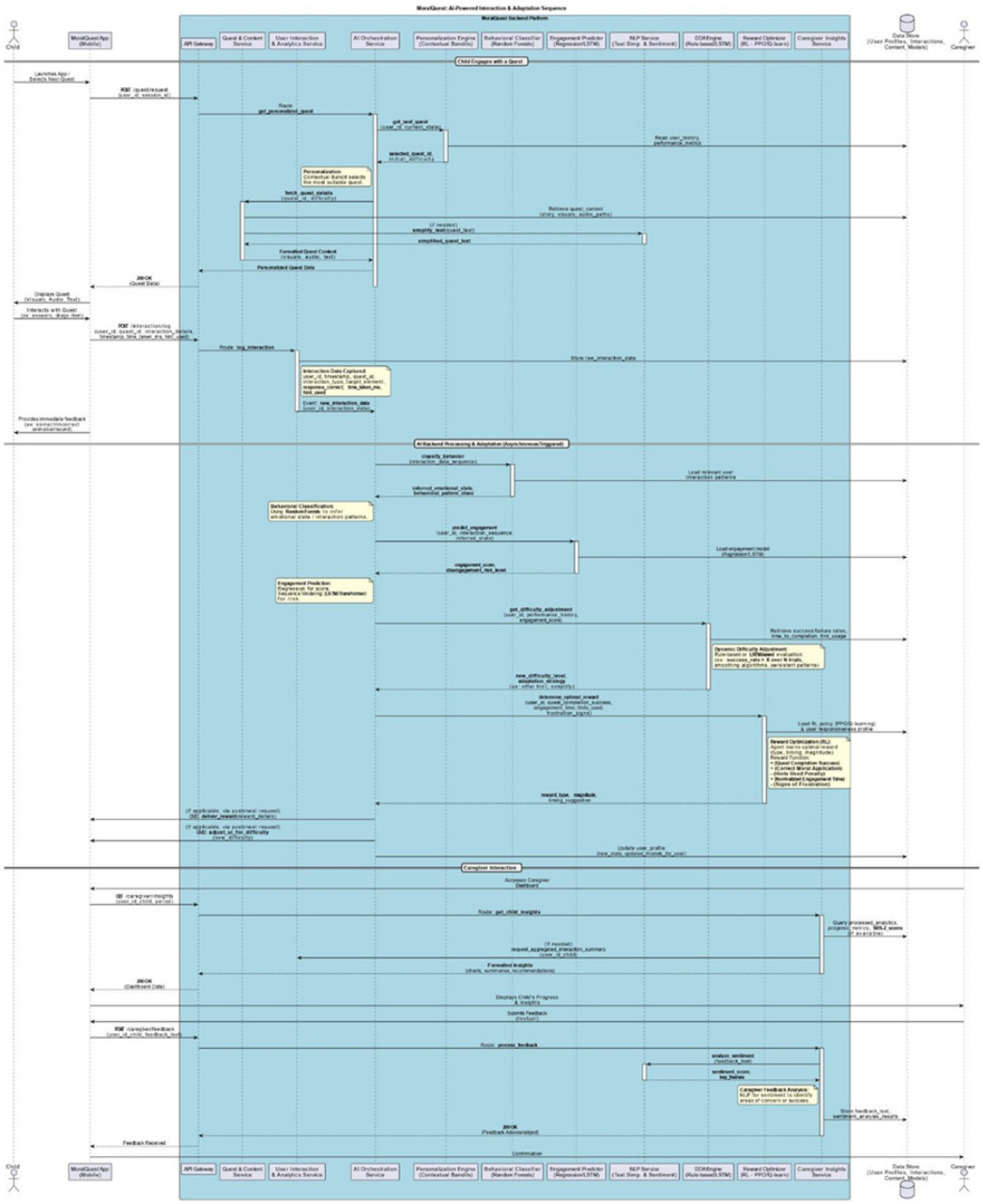


Figure 17.1: MoralQuest sequence diagram

## 5 References

[1] Hodges, H., Fealko, C., & Soares, N. (2020). Autism spectrum disorder: definition, epidemiology, causes, and clinical evaluation. Translational pediatrics, 9 (Suppl 1), S55–S65. Available online: https://doi.org/10.21037/tp.2019.09.09

[2] Tuna, A. (2023). Gamification as an Assistive Tool for Children With Autism Spectrum Disorder. In E. Efthymiou (Ed.), Inclusive Phygital Learning Approaches and Strategies for Students With Special Needs (pp. 151-168). IGI Global Scientific Publishing. Available online: https://doi.org/10.4018/978-1-6684-8504-0.ch007

[3] Kotsi, S.; Handrinou, S.; Iatraki, G.; Soulis, S.-G. A Review of Artificial Intelligence Interventions for Students with Autism Spectrum Disorder. Disabilities 2025, 5, 7. https://doi.org/10.3390/disabilities5010007

[4] Walsh O, Linehan C, Ryan C. Is there evidence that playing games promotes social skills training for autistic children and youth? Autism. 2025;29(2):329-343. doi:10.1177/13623613241277309

[5] Rahimi, M., Moradi, H., Vahabie, A.-. hossein ., and Kebriaei, H., "Continuous Reinforcement Learning-based Dynamic Difficulty Adjustment in a Visual Working Memory Game", arXiv e-prints, Art. no. arXiv:2308.12726, 2023. doi:10.48550/arXiv.2308.12726.

# Use Case 18: Smart Rehabilitation: AI-Driven Exercise Assessment and Live Feedback for Enhanced Recovery

**Organization**: Anna University, Chennai

**Country**: India

**Contact Person(s)**:

Dr. Dhananjay Kumar, dhananjay@annauniv.edu

Mr. B. Naren Karthikeyan, oraclenaren2004@gmail.com

Mr. A John Prabu, johnprabu0702@gmail.com

## 1 Use Case Summary Table

| Item | Details |
|---|---|
| Category | Healthcare |
| Problem Addressed | Patients with musculoskeletal disorders lack real-time corrective feedback and progress tracking when performing rehabilitation exercises at home, which can hinder recovery and lead to improper form or ineffective therapy [1]. |
| Key Aspects of Solution | • Real-time visual and audio cues to correct exercise through examining the extracted pose. Processing is to be done on a Raspberry Pi.<br>• Assessment of exercise to be done through a custom Spatio-Temporal Graph Neural Network (STGCN) on the cloud |
| Technology Keywords | STGCN, Raspberry Pi, Mediapipe, cross-platform Neural Network PACKage (XNNPACK) |
| Data Availability | Public (UI-PRMD Dataset: [7]) |
| Metadata (Type of Data) | Video data |
| Model Training and Fine-Tuning | • Live feedback is comparing joint angles and shifts in distance between the reference pose and the patient's pose.<br>• STGCN is to be trained using a triplet loss where the anchor is a correctly performed exercise, the positive sample is a correctly performed exercise other than the anchor, and the negative sample is an incorrect exercise.<br>• The final expected outcome is a latent space where correctly performed exercises are together and incorrect ones are further away from them. |

(continued)

| Item | Details |
|---|---|
| Testbeds or Pilot Deployments | We will use the facilities at the Media Research Laboratory of our university. Finally, we intend to test in a typical home environment. |
| Code repositories | STGCN: [8], Mediapipe: [9] |

# 2 Use Case Description

## 2.1 Description

Our solution is one that harnesses AI in assisting individuals suffering from musculoskeletal disorders, particularly in regions where there exists a significant shortage of physiotherapists. Therapeutic exercises are frequently prescribed as part of rehabilitation; however, in India, there is an average of merely 0.36 physiotherapists per 10,000 individuals, highlighting an urgent need for a digital solution. This use case presents a system designed to deliver real-time corrective feedback during physical therapy exercises [5]. The proposed system employs a simple camera and a Raspberry Pi as in Figure 1 utilising MediaPipe for pose detection and XNNPACK to accelerate inference. The system compares live user movements against reference joint angles and variations in distance between major key points [1]. The matching process is further refined by incorporating user-specific factors such as age and historical performance to ensure a personalised experience. Feedback is conveyed in textual, auditory, and visual formats via a display connected to the Raspberry Pi. Additionally, a bespoke Spatio-Temporal Graph Convolutional Network (STGCN) [2], hosted on the cloud, evaluates the performed exercises and generates an assessment score as in Figure 2. This score is displayed to assist users in monitoring their progress. Hence, it is safe to say our solution holds the potential to democratise access to expert guidance and corrective feedback, making at-home physiotherapy both more accessible and cost-effective.

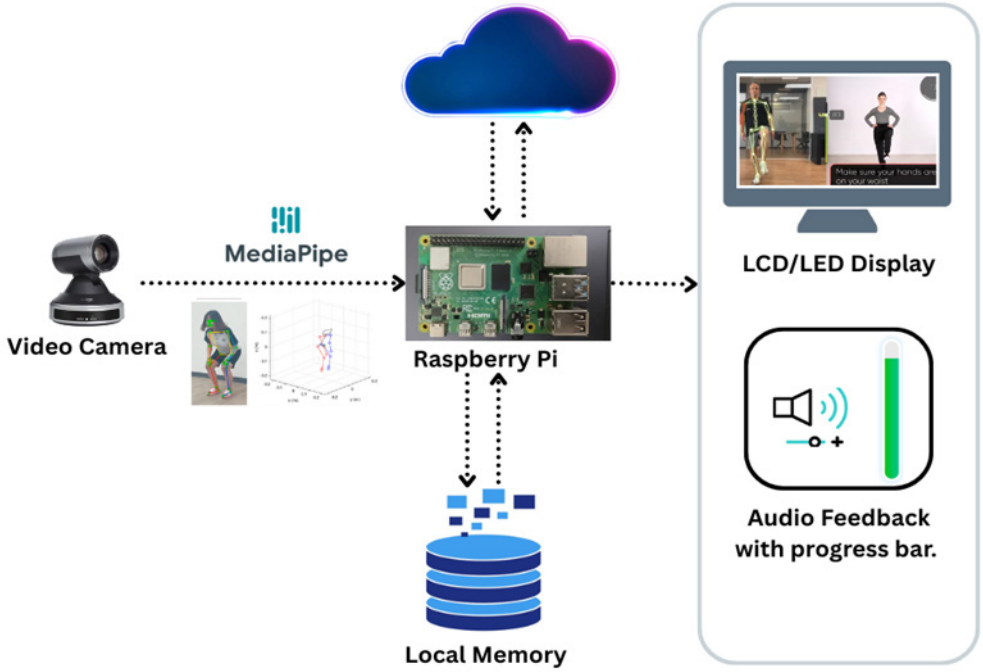



Figure 18.1: Overview of the proposed system

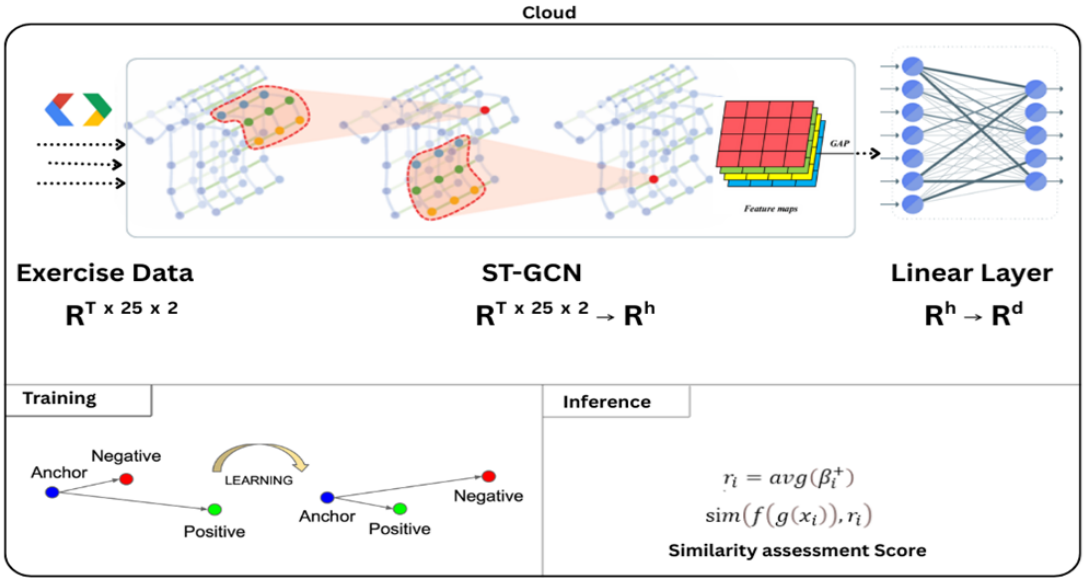



Figure 18.2: Schematics of the ST-GCN Model

**Use Case Status**: Ongoing–At inception stage.

**Partners**: We intend to get technical input and suggestions from the National Institute of Information and Communications Technology (NICT), Tokyo, Japan [6], and collaborate with them.

## 2.2 Benefits of use case

This system aims to reduce premature mortality from non-communicable diseases through prevention and treatment. By enabling real-time exercise assessment and adaptive rehabilitation, the system helps individuals recover from musculoskeletal disorders, post-stroke conditions, and post-surgical rehabilitation. It ensures users perform exercises correctly, reducing complications and improving health outcomes.

Traditional rehabilitation services are often inaccessible to rural populations and economically disadvantaged individuals due to high costs and limited healthcare infrastructure. By leveraging AI and automation, this system lowers the dependency on physical rehabilitation centres, making quality rehabilitation and exercise guidance available remotely through televisions, desktops, and mobile devices. This ensures that individuals, regardless of socioeconomic status or geographical location, can receive equitable healthcare support. By making rehabilitation cost-effective, scalable, and widely accessible, the system contributes to closing the health disparity gap and fostering inclusivity in healthcare

## 2.3 Future Work

- To overcome occlusion issues, future iterations will explore the integration of 3D sensing technologies, such as Microsoft Kinect. This will enable more accurate differentiation of complex movements (e.g., forward lean vs. backwards bend) and subtle movements that rely heavily on the trajectory rather than joint angles.
- Support for regional languages will also be extended, enhancing accessibility for non-English-speaking users.

# 3 Use Case Requirements

- **REQ 1**: It is critical that the system is capable of tracking and analysing human body movements in real time using a reliable pose estimation pipeline.
- **REQ 2**: It is critical that the system provides instant corrective feedback (audio and visual) by comparing user movement with reference data and ensures low-latency performance using optimised inference libraries such as XNNPACK.
- **REQ 3**: It is critical that the system is capable of utilizing an STGCN (Spatio-Temporal Graph Convolutional Network) model to assess exercise correctness accurately.
- **REQ 4:** It is expected that the system is maintained in a cloud-based repository of reference joint angles and movement patterns for comparison and dynamic updating.
- **REQ 5:** It is expected that the system will adapt feedback based on historical user performance, age, and progress trends to personalize the rehabilitation experience.
- **REQ 6:** It is expected that the system will adapt to new exercises by identifying consistent variations in joint angles and movement lengths over time.

# 4 Sequence Diagram

The sequence diagram in Figure 3 begins with the user performing an exercise after selecting the type of exercise. The data is passed through MediaPipe to extract poses, and then the reference data is compared with this to provide audio and visual cues. After completion of the exercise, the data is sent to the cloud, where the STGCN model gives the assessment score for the exercise. This performance report is sent to the user.

Figure 18.3: Sequence Diagram of our system

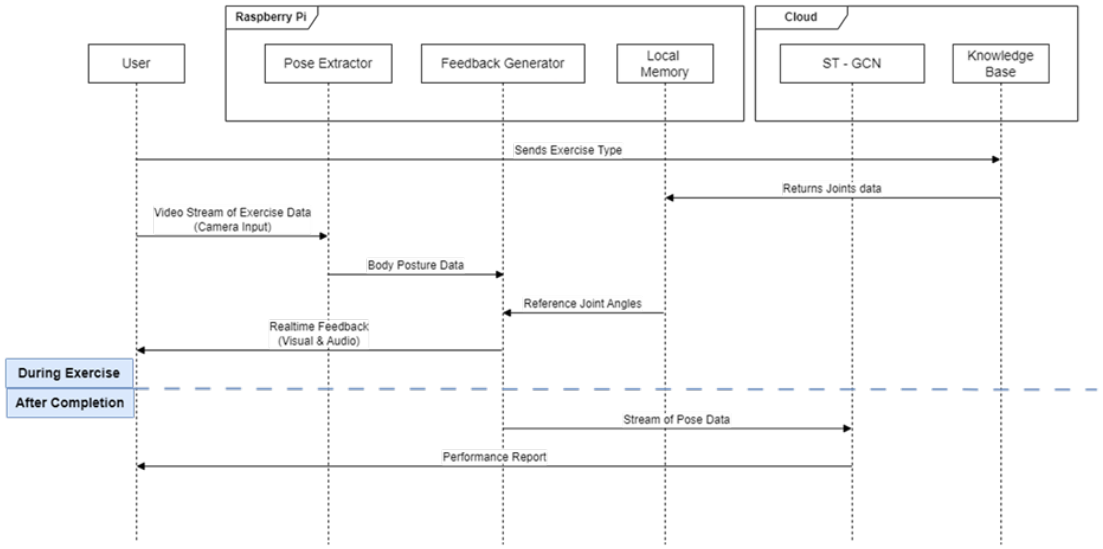



Figure 18.3: Sequence Diagram of our system

## 5 References

[1] H. Kotte, M. Kravčík, and N. Duong-Trung, "Real-Time Posture Correction in Gym Exercises: A Computer Vision-Based Approach for Performance Analysis, Error Classification and Feedback," in Proc. 3rd Int. Workshop on Multimodal Immersive Learning Systems (MILeS'23), 18th Eur. Conf. on Technology Enhanced Learning (EC-TEL 2023), Aveiro, Portugal

[2] G. Wei, H. Zhou, L. Zhang, and J. Wang, "Spatial–Temporal Self-Attention Enhanced Graph Convolutional Networks for Fitness Yoga Action Recognition," *Sensors*, vol. 23, no. 10, p. 4741, 2023

[3] S. Yan, Y. Xiong, and D. Lin, "Spatial Temporal Graph Convolutional Networks for Skeleton-Based Action Recognition," in Proc. AAAI Conf. Artif. Intell., vol. 32, no. 1, pp. 7444–7452, 2018.

[4] T. Chen, S. Kornblith, M. Norouzi, and G. Hinton, "A Simple Framework for Contrastive Learning of Visual Representations," in Proc. 37th Int. Conf. on Machine Learning (ICML 2020), PMLR, vol. 119, pp. 1597–1607, 2020

[5] Vakanski, H.-p. Jun, D. Paul, and R. Baker, "A Data Set of Human Body Movements for Physical Rehabilitation Exercises," Data, vol. 3, no. 1, p. 2, Jan. 2018

[6] National Institute of Information and Communications Technology (NICT), Japan. Available: https://www.nict.go.jp/en/.

[7] Open Data Lab, UI-PRMD Repository. Available: https://opendatalab.com/OpenDataLab/UI-PRMD/tree/main.

[8] hazdzz, STGCN GitHub Repository. Available: https://github.com/hazdzz/STGCN.

[9] Google AI Edge, MediaPipe GitHub Repository. Available: https://github.com/google-ai-edge/mediapipe.

# Use Case 19: AI Powered Webisite of UAE Government Healthcare System

EHS | Emirates Health Services United Arab Emirates

**Organization:** Emirates Health Services

**Country:** United Arab Emirates

**Contact person(s)**:

Mahmoud Alaa Omran, Mahmoud.omran@ehs.gov.ae, +971502616109,

Mrs. Amal Karmostaje, Amal.Karmostaje@ehs.gov.ae.

## 1 Use Case Summary Table

| Item | Details |
|---|---|
| **Category** | Healthcare |
| **Problem Addressed** | The primary challenge addressed by this use case is the inefficiency and inaccessibility of traditional healthcare service delivery across large public health systems. Patients often face long waiting times, complex registration processes, and fragmented service experiences, especially during peak hours or in high-demand facilities.<br>In addition, people with disabilities, language constraints, low digital literacy, or limited internet connectivity face disproportionate challenges in accessing digital health services. This limits the overall inclusiveness, efficiency, and satisfaction of public healthcare services. The public health system also lacked real-time adaptability, making it difficult to dynamically adjust operations based on live patient data, environmental factors, or user context (e.g., internet speed, location, device type).<br>This use case addresses these challenges by deploying a smart, AI-powered, patient-centric digital health platform that improves accessibility, efficiency, and service equity for all users. |
| **Key Aspects of Solution** | Streamlined healthcare delivery, enhanced patient engagement, AI chatbots, real-time dashboards, telemedicine solutions, AI-driven accessibility tools, multilingual AI support, reduced administrative costs, efficient data management, and remote care |

(continued)

| Item | Details |
|---|---|
| **Technology Keywords** | Generative AI, User-Centric Design, Mobile Health (mHealth), Digital Identity Verification, Regulatory Compliance (GDPR, HIPAA equivalents), Health Equity through AI, AI for Population Health Management, Smart Resource Allocation, AI for Health Crisis Response, Interoperable Health Systems, Data Integration Platforms, AI-Powered Automation, Predictive Modeling, Natural Language Processing (NLP), Personalized Healthcare, Remote Patient Monitoring, AI-Powered Chatbots, AI-Driven Diagnostics, Ethical AI in Healthcare, Fairness & Bias Mitigation, Healthcare Access Optimization, Cost-Efficiency in Health Services, Scalable Health Solutions, Real-Time personalization, Data-Driven Healthcare Information Access, Inclusive Innovation |
| **Data Availability** | 1. Public Data Sources:<br>• Aggregated health statistics are initially from Emirates Health Services, which serves as a national federal government healthcare services provider, and from other national government health authorities and international organizations (e.g., WHO, UAE Ministry of Health, Emirates Health Services open data portals) [4].<br>• Public Health Surveillance Data related to disease prevalence, vaccination rates, and population health indicators [1].<br>• De-identified Case Studies and Research Datasets are available through open-access medical or academic publications [6].<br>2. Private Data Sources<br>• Electronic Health Records (EHRs) from healthcare providers under EHS, used in compliance with data privacy regulations.<br>• Operational and Patient Flow Data from internal queue systems, self-check-in platforms, and UAE Pass-based digital interactions.<br>• Clinical and diagnostic data are gathered securely from telemedicine sessions and AI triage tools.<br>• Feedback and Survey Data collected through EHS mobile apps and service platforms are analyzed under ethical use policies. |
| **Metadata (Type of Data)** | Text, Numerical/Structured Data, Visual Data, Audio |

(continued)

| Item | Details |
|---|---|
| **Model Training and Fine-Tuning** | 1. Data Sources<br>The AI models are trained on a combination of<br>• Structured health system data (e.g., appointment logs, wait times, patient journey touchpoints)<br>• Anonymised user interaction data from mobile and web platforms<br>• Environmental and demographic data, including user location, internet speed, and device type<br>• Accessibility interaction logs to improve generative AI features for underserved populations<br>• All data is collected and processed in compliance with UAE data privacy laws and international health data regulations (HIPAA/ GDPR principles).<br>2. Initial Model Training<br>• Supervised learning is used for predictive models (e.g., queue/wait time estimation, appointment load balancing).<br>• Unsupervised & Semi-Supervised Learning support behavioral segmentation and clustering to optimize personalized digital health journeys.<br>• Large Language Models (LLMs) fine-tuned on healthcare-relevant Arabic and English corpora power Generative AI modules used in accessibility and patient assistance.<br>3. Fine-Tuning Approach<br>Continuous Learning Loops<br>• Models are periodically fine-tuned based on new patient flow patterns, updated scheduling policies, seasonal demand trends, and feedback from end-users.<br>Domain-Specific Adaptation<br>• Pre-trained base models are fine-tuned using health-specific vocabularies, clinical intent data, and real-world use cases from Emirates Health Services environments.<br>Personalization Layer<br>• Lightweight user context (e.g., prior visits, digital behavior, location, language preference) is injected into model responses for customized experience delivery.<br>Bias Auditing and Model Validation<br>• Regular audits ensure models perform equitably across different patient demographics, and validation datasets are refreshed to match evolving user behavior. |
| **Testbeds or Pilot Deployments** | It can be initiated by logging in to Emirates Health Services website: [1] |

# 2 Use Case Description

## 2.1 Description

The Emirates Health Services (EHS) has launched an AI-powered digital healthcare platform aimed at revolutionising healthcare accessibility and delivery across the Northern Emirates of the UAE. This innovative platform is designed to provide accessible, efficient, and inclusive healthcare services to all residents. Its primary objective is to guarantee universal healthcare access, streamline healthcare delivery, and significantly enhance patient engagement through the integration of cutting-edge technologies. These include AI chatbots, real-time dashboards, and advanced telemedicine solutions. The EHS platform directly tackles several long-standing issues prevalent in traditional healthcare systems. It seeks to overcome limited accessibility often faced by individuals with disabilities and those residing in rural areas, where infrastructural limitations can be a significant barrier. Furthermore, it aims to mitigate the high administrative costs and inefficiencies that stem from manual healthcare processes, which can lead to extended wait times and a notable absence of real-time data for patients seeking medical attention. The platform also addresses the critical issue of limited multilingual support, which frequently renders healthcare services inaccessible to a substantial number of expatriates and non-native speakers within the UAE. Unlike many current digital healthcare solutions, the EHS platform has been developed to specifically address common limitations. Many existing solutions often fall short in terms of inclusivity for individuals with disabilities and frequently provide insufficient support for users in rural areas who may have limited bandwidth. Additionally, a significant reliance on human resources in existing systems can lead to inherent inefficiencies and delays in service provision. The EHS platform leverages an AI-based approach to deliver a multitude of benefits. A 24/7 AI-powered chatbot, known as 'Ask EHS', offers instant assistance, thereby significantly reducing the workload on healthcare staff. The platform incorporates AI-driven accessibility tools that provide text-to-sign language translation, speech recognition, and adaptive user interfaces, ensuring that individuals with disabilities can navigate the system with ease. Real-time dashboards provide live data on crucial metrics such as hospital occupancy, appointment availability, and the current status of healthcare facilities, enabling faster and more informed decision-making. Telemedicine services are integrated to minimise the need for in-person hospital visits, which in turn saves both time and costs for patients. Crucially, multilingual AI support ensures that the platform's services are accessible in multiple languages, catering to the diverse population of the UAE. Despite its numerous advantages, the AI-based approach does present some potential drawbacks. The initial implementation costs for establishing the necessary AI infrastructure can be substantial. Moreover, a certain dependence on internet connectivity may pose challenges for some users; hence, a lightweight version of the platform has been developed to mitigate this issue [5].

**Use-case status:** The use case is part of a larger product development

### Partners:

1 UAE Pass: National digital identity platform integrated into EHS systems for secure authentication, appointment confirmation, and self-check-in [1].

2 Telecommunications and Digital Government Regulatory Authority (TDRA): Supporter of the Zero Government Bureaucracy initiative and enabler of UAE Pass integration across government services [2].

3 Ministry of Health and Prevention—UAE (MOHAP UAE) [3].

4 Collaborative partner in aligning health innovation with national digital health strategy and policy frameworks.

5 Federal Authority For Identity, Citizenship, Customs & Port Security (ICP UAE) [6]

## 2.2 Benefits of use case

- EHS.gov.ae enhances healthcare accessibility through telemedicine, real-time dashboards, and AI-driven services, ensuring universal health coverage for all, including People of Determination and rural populations.Promotes preventive care, mental health support, and maternal health services through AI-powered consultations and digital resources.
- Supports continuous medical education for healthcare professionals via the Maharati e-learning platform and annual conferences such as the Nurse and Mental Health Conferences.
- Enhances public health literacy through AI-generated medical illustrations, blogs, and open data resources.
- Provides accessible digital healthcare services for women, including maternal and reproductive health support, through telemedicine and AI-driven platforms. Integrates AI, IoT, and cloud-based solutions to deliver innovative digital healthcare services, optimizing hospital resource management and patient care.
- Ensures equal access to healthcare through AI-powered accessibility tools, multilingual support in 120+ languages, and a lightweight website version for low-bandwidth users.
- Promotes sustainable healthcare practices by reducing paper use through digital medical records, e-prescriptions, and AI-driven hospital management systems.
- Reduces carbon emissions by encouraging telemedicine, minimizing travel for medical services, and optimizing hospital energy use through AI-driven resource management.

## 2.3 Future works:

In the future iteration of this product, we will be working on

- Advanced AI Triage & Risk Prediction Models
  EHS plans to develop AI models that analyze patient symptoms and historical data to suggest triage levels or flag high-risk cases in real time. This would help optimize clinical resources and reduce emergency room overcrowding.
- Expansion of Generative AI for Accessibility
  Broaden the use of generative AI to include real-time translation, voice-guided navigation, and AI-generated simplified medical explanations, ensuring equitable service for users with disabilities, low digital literacy, or language barriers.
- Integration with Wearables and IoMT
  Introduce data streams from wearable health devices and Internet of Medical Things (IoMT) sensors to enable continuous remote monitoring and preventive care recommendations.
- Personalized Digital Health Journeys
  Using AI and behavior analytics, tailor health information, reminders, and services based on user history, preferences, and medical needs, improving patient engagement and outcomes.

# 3 Use case Requirements

- **REQ-01:**It is critical that the system integrate with UAE Pass for secure authentication, user verification, and seamless access to healthcare services.

- **REQ-02:** It is critical that the AI models operate on real-time or near-real-time data from various sources (EHRs, queue systems, mobile apps) to support timely decision-making.
- **REQ-03:** It is critical that the solution be compatible with Emirates Health Services' digital infrastructure, including health information systems and patient management platforms.
- **REQ-04:** It is critical that the AI model is scalable across multiple healthcare facilities, with the ability to process increasing volumes of data and serve a diverse patient population.
- **REQ-05:** It is critical that all data handling must comply with national and international healthcare data regulations (e.g., UAE data protection laws, HIPAA/GDPR principles).
- **REQ-06:** It is critical that the platform support Arabic and English to ensure accessibility and inclusiveness for all users.
- **REQ-07:** It is critical that the mobile and web interfaces is intuitive and responsive to allow easy appointment booking, check-in, and service tracking by patients.
- **REQ-08:** It is critical that the system use AI to predict peak times, optimize queue management, and recommend operational adjustments in real-time.
- **REQ-09:**It is critical that the platform generate performance reports and operational insights for decision-makers to support data-driven planning.
- **REQ-10:** It is expected that the platform include fallback mechanisms to ensure service continuity in case of connectivity issues or technical failures.
- **REQ-11:** It is critical that the web solution utilize AI techniques (e.g., adaptive loading, content prioritization, compression) to ensure fast loading times and minimal bandwidth use, especially in low-connectivity environments.
- **REQ-12:** It is critical that the system use generative AI (e.g., voice-to-text, text simplification, automatic summarization, audio guides) to assist users with disabilities, low literacy, or language barriers in navigating digital services effectively.

## 4 Sequence Diagram

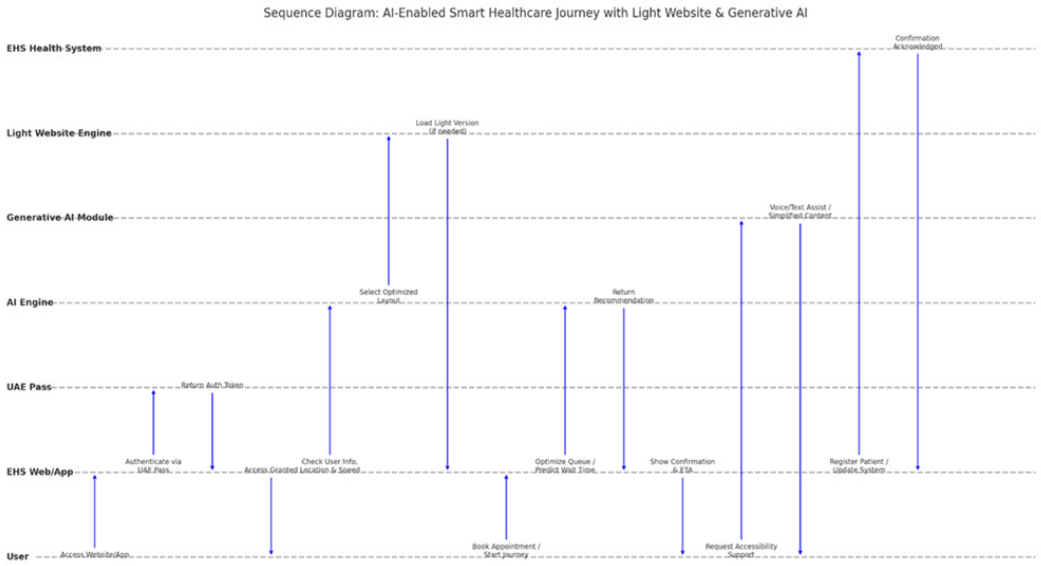



Figure 19.1: Sequence Diagram of our system

## 5 References

[1] Emirates Health Services Annual Report Includes documentation on digital initiatives, smart patient management systems, and UAE Pass integration. https://ehs.gov.ae

[2] UAE Digital Government Strategy 2025,Telecommunications and Digital Government Regulatory Authority (TDRA), https://tdra.gov.ae

[3] Zero Government Bureaucracy Program– UAE Cabinet, https://u.ae/en/about-the-uae/digital-uae

[4] World Health Organization – Global Strategy on Digital Health 2020–2025, https://www.who.int/publications/i/item/9789240020924

[5] UAE Artificial Intelligence Strategy 2031 Ministry of Cabinet Affairs and The Future, https://ai.gov.ae

[6] Federal Authority For Identity, Citizenship, Customs & Port Security (ICP UAE), https://icp.gov.ae/en/

## Use Case 20: An AI-Powered Augmentative and Alternative Communication (AAC) tool

**Organization**: AI4Africa

**Country**: Nigeria

**Contact Person(s)**:

Emmanuella Sule suleemmanuella0010@gmail.com

Dr Houda Chihi houda.chihi@supcom.tn

Emmanuel Aaron aaronemmanuel054@gmail.com

Chidi Ebube chidizack24@gmail.com

Emmanuel Ani ani.mlengineer@outlook.com

Okafor Miracle Uche okaformiracle212@gmail.com

Blessed Guda gudablessed@gmail.com

### 1 Use Case Summary Table

| Item | Details |
|---|---|
| Category | Healthcare |
| Problem to be addressed | Challenges in communications for speech-impaired individuals. |
| Key solution | Conversion and reconstruction of unintelligible dysarthria speeches. |
| Technology keywords | Communication disorder, Artificial Intelligence, Automatic Speech Recognition system, speech reconstruction |
| Data availability | Private. |
| Metadata (type of data) | Structured and Unstructured (speaker ID, age, audio recording). |
| Model Training and Fine-Tuning | Supervised pre-training, transfer knowledge distillation(KD), fine-tuning on a small dataset |
| Testbeds or pilot Deployments | Limi Hospital, Nigeria* |

(continued)

| Item | Details |
|---|---|
| Code repositories | Nill |

# 2 Use Case Description

## 2.1 Description

A speech disorder, also referred to as speech impairment or speech impediment, is a condition that affects an individual's ability to produce the sounds necessary to form words effectively [1]. This often manifests as difficulties in articulation, where speech sounds are substituted, omitted, added, or distorted. Speech disorder is a common condition with individual dysarthria, non-verbal autism, or individuals with neurodevelopmental issues in general.

Speech is a fundamental means of expressing thoughts and emotions. However, individuals with speech disorders face communication challenges that can negatively affect their self-esteem and overall quality of life. Various support strategies exist to address these difficulties. Traditional therapy approaches may include sign language, symbol systems, or communication boards. Recent advancements in technology have introduced Assistive Technology (AT), such as Augmentative and Alternative Communication (AAC) tools, which offer support to enhance communication. AAC tools are fundamental in helping individuals with speech disorders overcome communication challenges, offering alternative ways to express needs and desires. They encompass a wide range of assistive strategies and devices, such as an application that allows users to build sentences by tapping icons. AAC supports learning, helps develop language and cognitive skills, and improves societal inclusion by enabling individual participation.

This use case proposes an AI-powered AAC tool that can turn unintelligible speech with speech disabilities into intelligible and clear speech to enhance their communication independence. While traditional AAC tools provide crucial support, the integration of Artificial Intelligence (AI) into AAC and other assistive technologies is an emerging field with substantial potential to offer more personalized, adaptive, and effective interventions for individuals with speech disorders, and be incorporated into their wider health therapy [2].

## 2.2 Benefits of the use case

This AI-powered AAC device provides individuals with neurodevelopmental conditions an effective means of communication. The inability to express thoughts and emotions often leads to frustration, which can escalate into anxiety and stress. By enabling individuals to communicate their needs and emotions, this technology alleviates frustration, improves mental well-being, and enhances interpersonal relationships, ultimately leading to a better quality of life.

Moreover, this AI-powered AAC device promotes social inclusion. Communication barriers often result in social isolation for individuals, limiting their participation in everyday interactions. By empowering them with a reliable means of expression, this tool fosters engagement in conversations, promotes social interaction, and helps reduce biases and inequalities they may face.

Additionally, the continuous use of this assistive AI tool strengthens learning and communication skills. Beyond serving as a communication aid, it functions as an AI-assisted therapy tool, complementing multidisciplinary treatment approaches. By encouraging consistent engagement and practice, it enhances cognitive and linguistic abilities, aiding personal development and educational growth.

## 2.3 Future Works

This project aims to build an AI-powered AAC device to address communication challenges for individuals with speech disability such as dysarthria.

The proposed methodology outlines a comprehensive approach leveraging speech-encoder and text-encoder for an end-to-end voice conversion system [3]

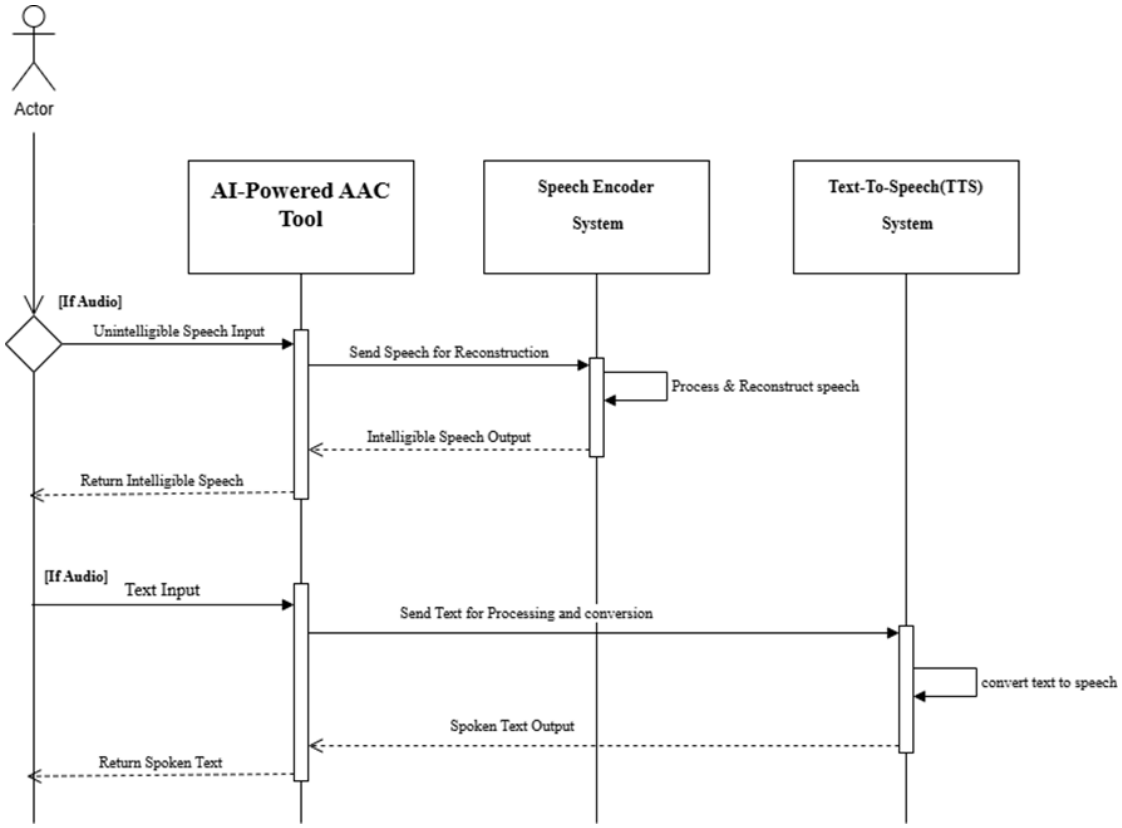
1. TTS System Training: A seq2seq Text-to-Speech (TTS) system, specifically based on the Tacotron architecture, is first trained using transcribed normal speech. This system includes an encoder, attention mechanism, and decoder.
2. Cross-Modal KD for Speech-Encoder Training: A teacher-student framework is employed for KD. The text encoder from the pre-trained TTS system acts as the "teacher". A separate speech encoder (the "student") is trained using transcribed dysarthric speech. The goal is to train the speech encoder to generate linguistic representations (spectral embeddings) from dysarthric speech that are similar to the representations (character embeddings) produced by the text encoder from text.
3. End-to-end voice construction: The trained speech encoder replaces the text encoder and is concatenated with the attention and decoder modules from the original TTS system. This forms the final end-to-end voice system. During conversion, this system takes the dysarthric speech's spectral features as input, generates spectral embeddings via the speech encoder, and then uses the TTS attention and decoder modules to predict a normal mel-spectrogram. A WaveRNN vocoder synthesizes the final waveform.
4. Dataset: The UASpeech dataset will be used for the cross-modal knowledge distillation phase to train the speech encoder. The UASpeech dataset is a speech dataset for dysarthric speech (i.e., speech impaired by motor disorders) developed by the University of Illinois at Urbana-Champaign (UIUC). It contains speech from 15 dysarthric speakers and 13 control (non-dysarthric) speakers.
5. Fine-tuning: speech-text dataset of dysarthric speeches or people with speech disability will be collected from hospitals in Nigeria, with affiliation will be pursued. This dataset will be used to fine-tune the pre-trained voice conversion model.
6. Affiliations with Limi Hospital, Nigeria, or other hospitals in Nigeria will be actively pursued in order to test and deploy the proposed solution.

# 3 Use Case Requirements

- **REQ-01:** It is critical that the system implements an end-to-end voice conversion model capable of reconstructing unintelligible speech.
- **REQ-02:** It is critical that the system features multimodal input, that is, audio and text.
- **REQ-03:** It is critical that the system trains on dysarthria or non-standard speech data for an effective speech conversion and reconstruction that will typically fail in an ASR trained on standard data.
- **REQ-04:** It is critical that the system implements data privacy and encryption measures to protect the sensitivity of the data.

- **REQ-04:** It is critical that the system finds a balance between generalization and personalization for its users.

## 4 Sequence Diagram




## 5 Reference

[1] Emirates Health Services Annual Report, Includes documentation on digital initiatives, smart patient management systems, and UAE Pass integration. https://ehs.gov.ae

[2] UAE Digital Government Strategy 2025,Telecommunications and Digital Government Regulatory Authority (TDRA), https://tdra.gov.ae

[3] Zero Government Bureaucracy Program– UAE Cabinet, https://u.ae/en/about-the-uae/digital-uae

[4] World Health Organization – Global Strategy on Digital Health 2020–2025, https://www.who.int/publications/i/item/9789240020924

[5] UAE Artificial Intelligence Strategy 2031 Ministry of Cabinet Affairs and The Future, https://ai.gov.ae

[6] Federal Authority For Identity, Citizenship, Customs & Port Security (ICP UAE), https://icp.gov.ae/en/

# Use Case 21: AI-Based Health Monitoring and Advisory System

**Organization:** Mbeya University of Science and Technology (MUST), Department of Electronics and Telecommunication Engineering

**Country:** Tanzania

**Contact Person**:

Teophiler Daud Mage maggetheo@gmail.com

## 1 Use Case Summary Table

| Item | Details |
|---|---|
| Category | Healthcare, Intelligent transport, accessibility |
| Problem addressed | Limited access to health care and insufficient awareness and actionable advice for maintaining healthy lifestyles |
| Solution to the problem | Integrating AI with wearable sensors for real-time alerts and automated health advice generation |
| Technology keywords | Artificial Intelligence, Monitoring |
| Metadata (type of data) | |
| Data availability | Sensor-based data from wearable devices |
| Model training | Simple machine learning models and anomaly detection |
| Testbeds or pilot deployments | |

## 2 Use Case Description

### 2.1 Description

In developing countries access to regular medical checkups is limited, especially in rural areas. Non-communicable diseases (NCD) such as hypertension and diabetes often go undetected due to the absence of early warning systems. This use case introduces an AI-Based Health Monitoring and Advisory System using wearable sensors and wireless communication. The system collects vital health parameters like heart rate, $SpO_2$ levels, and body temperature, processes them using AI models, and issues real-time alerts and advisory messages when abnormal values are detected.

The innovation lies in its low-cost, portable, and Artificial Intelligence (AI)-powered solution that enables continuous health monitoring without requiring patients to visit healthcare centres. The expected impact includes reduced NCD-related complications, increased health awareness, and improved healthcare accessibility.

### Partners

- Mbeya University of Science and Technology (MUST)
- Local Health centres

## 2.2 Benefits of the use case

- The system promotes early diagnosis and preventive care.
- Encourages digital health innovations and ICT use in healthcare.
- Targets healthcare access disparities between urban and rural populations.

The project contributes to improving NCD prevention, healthcare accessibility, and innovation in health systems. For instance, the system promotes early diagnosis and preventive care; encourages digital health innovations and ICT use in healthcare, and targets healthcare access disparities between urban and rural populations.

## 2.3 Future Work

Future development of the system includes integrating more biosensors, such as Electrocardiogram (ECG), respiratory rate, and glucose monitors, to provide a comprehensive health profile. Advanced AI models, such as deep learning algorithms, will be implemented to improve prediction accuracy and personalization.

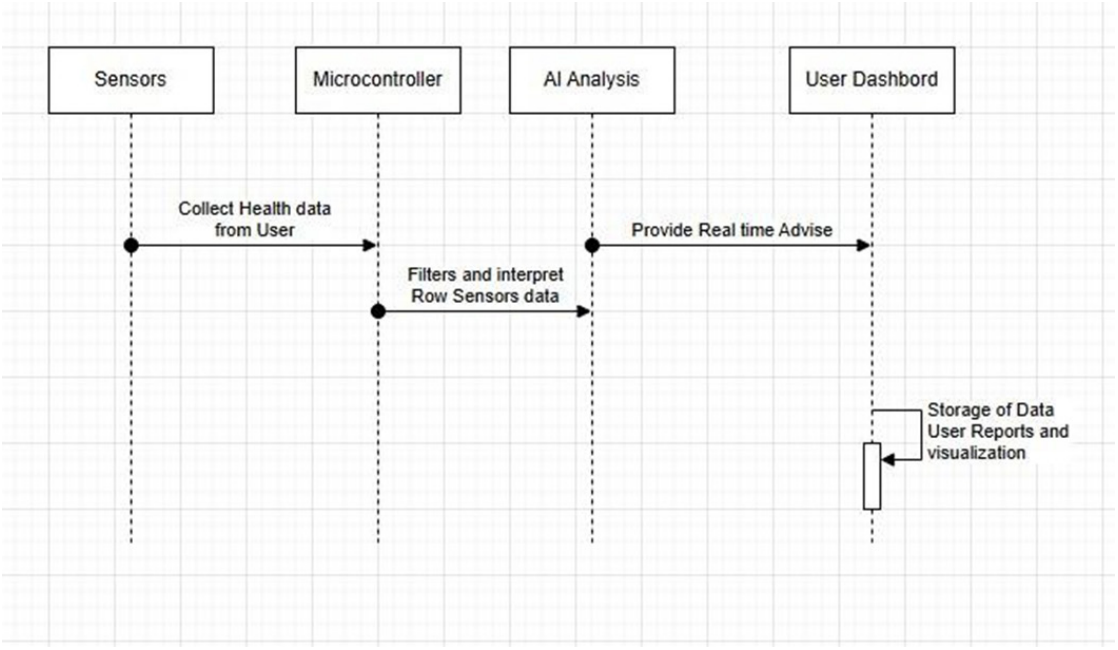
Collaboration with telecom providers can enhance real-time data transmission in rural areas. Additionally, a cloud-based dashboard for healthcare professionals will be developed to allow remote consultations.

# 3 Use Case Requirements

- **REQ-01:** Real-time data acquisition from wearable sensors
- **REQ-02:** Wireless transmission of health data
- **REQ-03:** On-device AI/ML capability for anomaly detection
- **REQ-04:** Secure storage and privacy protection for user data
- **REQ-05:** Mobile/web interface for alerts and health insights
- **REQ-06:** Light Emitting Diode (LED) indication for normal/abnormal condition on the device
- **REQ-07:** Scalability to support multiple users in rural settings

# 4 Sequence Diagram

This diagram shows the flow of data and interactions between sensors, AI processing, alert generation, and healthcare access.




## 5 References

[1] World Health Organization, "Non communicable diseases," 2022. Available: https://www.who.int/news-room/fact-sheets/detail/noncommunicable-diseases. [Accessed: May 25, 2025]

[2] "Pulse Sensor Arduino Guide," PulseSensor.com. Available: https://pulsesensor.com/.

[3] Arduino, "Arduino Documentation." Available: https://www.arduino.cc/.

[4] United Nations, "SDG reports." Available: https://sdgs.un.org/goals.

# Use Case 22: EquiDermAI: An Inclusive Deep-Generative Framework to Overcome Racial & Regional Barriers in Skin Disease Diagnosis

**Organization:** Indian Institute of Technology, Madras

**Country:** United States

**Contact Person(s)**:

Akshat Santhana Gopalan, akshatsg@outlook.com

Narayanan Madaboosi Srinivasan, narayananms@smail.iitm.ac.in

## 1 Use case summary table

| Item | Detail |
|---|---|
| Category | Healthcare |
| The problem to be addressed | In today's world, AI and computer vision-based diagnosis of dermatological diseases are playing a key role in the healthcare industry. With a rapidly increasing number of medical institutions relying on such intelligent models for skin disease diagnosis, it is vital to ensure their accuracy is maintained across different populations.<br>However, a pivotal problem arises where a bias in the training data in terms of the skin colour of the training instances prevents consistent accuracy across all skin tones. Most dermatological training datasets are heavily skewed to contain a majority of white skin tones, and models trained on such datasets perform poorly on diagnosing other darker, underrepresented skin tones. Black patients with psoriasis receive advanced treatments less often (8.3% vs. 13.3% for white patients). Melanoma survival rates starkly differ: 93% for white adults, but only 71% for Black individuals (American Cancer Society, 2023). Research on top skin conditions for people of colour is alarmingly scarce (Takeshita et al). The digital divide exacerbates these disparities.<br>Many underdeveloped regions in the world do not have access to powerful diagnostic tools. About 32% of the global population was said to have no internet access in 2024, displaying a clear digital divide and inaccessibility to such diagnostic tools (International Telecommunication Union, 2020). The current scenario thus has prominent racial and digital divisions around the world in the field of dermatological healthcare. |

(continued)

| Item | Detail |
|---|---|
| Key aspects of the solution | The proposed solution to these major problems is **EquiDermAI**, an intelligent framework for overcoming racial and digital barriers in AI diagnosis of any skin-related lesion/disease.<br><br>There are 2 main components that address the two major problems previously explained:<br><br>1. **Deep-Generative Style Transfer for Skin Tone Data Augmentation:** This component addresses the racial divide in intelligent dermatological diagnosis. EquiDermAI uses an image-to-image translation method based on CycleGAN (Zhu et al., 2017) to perform semantic-preserving transformation of skin lesion images. Specifically, it adapts lesion images from overrepresented lighter skin tones to reflect underrepresented darker skin tones while preserving lesion morphology and clinical features. Unlike classical neural style transfer, which may distort structural pathology, CycleGAN operates without requiring paired training data and focuses on domain adaptation, enabling realistic generation of lesions on diverse skin tones without altering diagnostic-relevant morphology. By augmenting datasets with these clinically faithful synthetic images, EquiDermAI improves diagnostic model generalisation across skin types and helps reduce misdiagnosis rates in underrepresented populations.<br>2. **Quantized TinyML Development for Resource-Constrained Areas:** This component addresses the digital divide. With many regions in the world lacking access to powerful technological tools, it is important to keep diagnostic resources very efficient. EquiDermAI further optimizes AI models for low-power devices, enabling accurate diagnosis in resource-constrained areas with limited computing power or internet connectivity (Warden & Situnayake, 2019).<br><br>These two components are intricately designed to contribute to the overall goal of creating an inclusive and accessible skin disease diagnosis system. By combining advanced deep learning techniques with efficient deployment strategies, EquiDermAI aims to overcome both racial and regional barriers in dermatological care, democratizing accurate skin disease diagnosis.<br><br>**CONSIDERATIONS**: There are legitimate concerns regarding whether generated images retain diagnostic fidelity, particularly when transforming skin tone while preserving the clinical morphology of lesions. To address this, EquiDermAI will involve expert validation from board-certified dermatologists and skin-of-colour specialists in future iterations. These experts will assess the realism and clinical accuracy of the synthetic images, and validation may incorporate structured annotation protocols with inter-rater reliability metrics (e.g., Cohen's kappa) to quantify agreement. Alternatively, or in parallel, automated similarity assessments may be employed by comparing the learned feature embeddings of real and generated lesions using pre-trained dermatological classification networks. Techniques such as perceptual loss or cosine similarity in latent space can serve as quantitative proxies for evaluating whether lesion morphology remains preserved during style transformation. These safeguards aim to ensure that synthetic data used for training maintains clinical utility and does not introduce misleading artifacts. |
| Technology keywords | Deep Learning, Generative Adversarial Network (GAN), Data Augmentation, tinyML, TensorFlow Lite, Computer Vision, Swift/CoreML |
| Data availability | Private |

(continued)

| Item | Detail |
|---|---|
| Metadata (type of data) | Images and Labels |
| Model Training and Fine-Tuning | Generative Adversarial Networks (GANs), Convolutional Neural Networks (CNNs) |
| Testbeds or Pilot Deployments | Initial prototyping efforts are focused on optimising the model for deployment on resource-constrained mobile hardware, such as entry-level smartphones powered by Acorn RISC Machine (ARM) Cortex-A series chipsets with limited Random Access Memory (RAM) (<2 GB) and no reliance on cloud connectivity. While full validation is ongoing, the quantized model has been designed with considerations for low latency and a compact memory footprint suitable for on-device inference. Future benchmarking will evaluate inference speed, energy efficiency, and diagnostic accuracy across a range of low-power devices to ensure equitable real-world applicability. |
| Code repositories | Not publicly disclosed |

# 2 Use case description

## 2.1 Description

This use case aims to democratise smart dermatological diagnosis by tackling both racial and digital divisions. Many diseases are increasingly being diagnosed by AI and computer vision. However, there is a significant amount of difficulty and inaccuracy in the diagnosis of darker skin due to the bias in training datasets - the majority of training instances are based on white skin lesions. This makes darker skin tones more underrepresented in AI-based dermatological diagnosis systems, causing increased misdiagnosis rates for such populations. Another major problem is the digital divide; even with accurate diagnostic tools, many regions in the world lack the technology needed to run high-power tools. The combination of these two problems represents a major disparity in dermatology. A common question is, why not just remove the colour from all images to ensure equality? This is because colour plays a pivotal role in skin diseases and provides very valuable information and patterns to the AI model. This use case delves into the solution EquiDermAI, a deep-generative framework that addresses both racial divisions and the regional inaccessibility to these diagnostic tools (the digital divide). For countless diseases, many datasets possess a much larger proportion of lighter skin tones, and the AI models trained on such datasets may classify lighter skin tones accurately but misdiagnose the darker, more underrepresented skin tones. EquiDermAI, for any given disease, takes as input a dataset, distinguishes the lighter skin tones (more populous in the dataset) from the darker ones, and, using the few darker skin tone training instances as reference, carries out a generative style transfer process using the powerful generative adversarial networks algorithm (GANs) to generate novel training instances of multiple different skin tones of the lesion from the large majority of lighter skin tones. The style transfer process continually learns to refine the pigmentation of the lesions as well, and thus, can generate new images of the white skin lesions transferred to other darker underrepresented skin tones that accurately depict the different ways a lesion would look across different skin tones. The datasets used are expert-validated, and thus the reference images of the darker skin tones provide a measure to ensure that the

generated images of lesions on darker skin tones are realistic and imitate how a true lesion on that skin would look.

This approach could not only improve the accuracy of lesion diagnosis on underrepresented skin tones but also serve as an efficient, proactive diagnostic tool to help prevent disease spread. For example, during the Mpox outbreak, which was largely concentrated in the Congo and surrounding regions of Africa, with only a few cases elsewhere (World Health Organization, 2022), the availability of images showing Mpox on darker skin tones could be used to generate synthetic images for lighter skin tones. This could significantly enhance model accuracy and offer a preventive diagnostic tool for identifying new cases of Mpox on lighter skin tones. This example highlights that, for certain diseases, the need for data might even require generating conversions from darker to lighter skin tones instead, as described above. Once the datasets are produced/augmented, anyone can train different classification models on them. However, most of the models are not optimized to run efficiently. EquiDermAI goes a step further to train the model on the datasets produced (now with the skin tones equally represented) and quantizes them in order to run efficiently with tinyML. This process ensures that the model can be run on low-power devices with little storage and no internet connection, thus being more accessible to doctors and other individuals in underdeveloped regions with resource constraints to ensure proper diagnosis. Overall, EquiDermAI aims to address as many diseases with discriminatory training biases as possible in order for everyone to receive accurate diagnoses. It also aims to provide everyone, regardless of where they are in the world, with the proper efficient AI-driven diagnostic solutions in order to ensure an equitable future in dermatology

**Use Case Status**: Ongoing–In active deployment and continuous development.

**Partner(s):** None

## 2.2 Benefits of use case

EquiDermAI improves the early and accurate diagnosis of diseases across diverse populations, and this directly contributes to enhanced health outcomes and reduced mortalities, further enabling such diagnostic tools even on low-power devices for those in resource-constrained areas, reaching as many people as possible, supporting universal access to crucial healthcare services.

This framework mitigates both racial and regional biases in the field of dermatological diagnosis with AI, thus ensuring fair diagnostic performance across diverse skin tones and underserved populations. This promotes equality in the delivery of such pivotal healthcare services, addressing disparities in the access and quality of healthcare.

By leveraging tinyML and edge computing, EquiDermAI minimizes computational resources and avoids reliance on cloud-based infrastructure, reducing energy consumption. It fosters sustainable AI deployment, particularly in regions with limited digital infrastructure.

## 2.3 Future work

Currently, EquiDermAI is able to generate its own skin lesion training instances with underrepresented skin tones that accurately mimic that of expert-validated images and training models that can use these datasets and run on low-power devices to perform accurate classifications. However, future works will focus on expanding EquiDermAI as an app for several

diseases and those that are in the state of outbreaks to provide everyone with a plausible solution to diagnosis, regardless of race or region. By proactively creating skin-tone-inclusive diagnostic frameworks for underrepresented and data-biased conditions, EquiDermAI aims to serve as a critical resource for global health equity and preparedness.

## 3 Use case requirements

- **REQ-01**: It is critical that the diagnostic framework integrates deep-generative augmentation techniques to expand datasets with diverse skin tones synthetically, ensuring balanced model performance across all racial and ethnic groups.
- **REQ-02**: It is critical that the system supports lightweight, quantised AI models optimised for deployment on low-power edge devices to enable accessibility in remote and resource-constrained regions.
- **REQ-03**: It is critical that the system include tools to simulate how emerging diseases (e.g., monkeypox, measles) present across different skin tones, supporting rapid, inclusive diagnosis during public health emergencies.
- **REQ-04**: It is critical that the framework ensures privacy by performing all inference locally on-device, with no data transmission or centralised aggregation, thereby containing sensitive health data entirely within the user's hardware and mitigating exposure risks without requiring encryption or differential privacy layers.
- **REQ-05**: It is expected that all model training and evaluation will incorporate fairness metrics—including subgroup Receiver Operating Characteristic-Area Under the Curve (ROC-AUC), demographic parity difference, and equality of opportunity, with deployment thresholds enforcing less than 5% performance deviation across stratified skin tone groups.

## 4 Sequence diagram

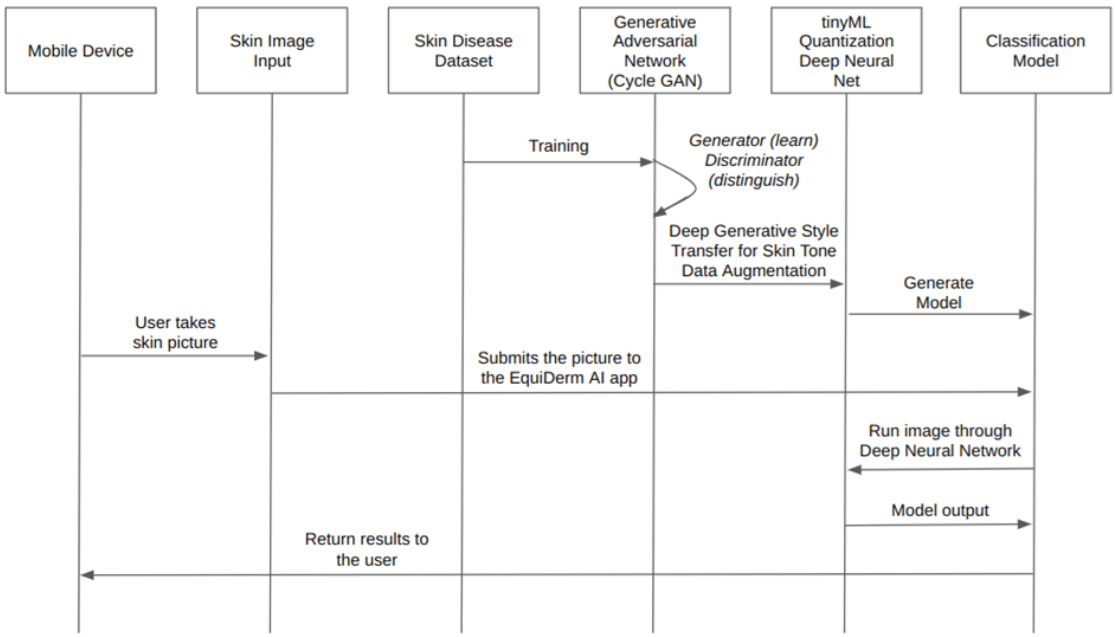



Figure 22.1: Our System sequence diagram

## 5 References

[1] Takeshita J, Gelfand JM, Li P, et al. Psoriasis in US Adults: Prevalence Estimates from Medical Expenditure Panel Surveys, 2009-2011. JAMA Dermatol. 2021;157(8):940-946

[2] American Cancer Society. Cancer Facts & Figures 2023. Atlanta: American Cancer Society; 2023

[3] International Telecommunication Union. Measuring Digital Development: Facts and Figures 2020. Geneva: ITU Publications; 2020

[4] Goodfellow I, Pouget-Abadie J, Mirza M, et al. Generative Adversarial Nets. Advances in Neural Information Processing Systems. 2014;27

[5] Zhu JY, Park T, Isola P, Efros AA. Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks. In: Proceedings of the IEEE International Conference on Computer Vision. 2017:2223-2232

[6] Warden P, Situnayake D. TinyML: Machine Learning with TensorFlow Lite on Arduino and Ultra-Low-Power Microcontrollers. O'Reilly Media; 2019

[7] World Health Organization. Multi-country outbreak of mpox, external situation report #10 - 14 September 2022. Geneva: WHO; 2022

# Use Case 23: Empowering Student Mental Well-being with AI: Intelligent Journaling, Emotional Insights, and Personalized Growth

**StudentSphere** - Fostering peace, nurturing minds

**Organization:** Vellore Institute of Technology, Chennai, India

**Country**: India

**Contact Person(s)**:

Mohammed Sayeed A, mohammedsayeed.a1@gmail.com

Bushra Sabreen, bushra.sabreen02@gmail.com

Dr Suganya G, suganya.g@vit.ac.in

## 1 Use case summary table

| Items | Details |
|---|---|
| Category | Healthcare |
| Problem Addressed | Students across India are experiencing a significant increase in mental health challenges, yet they lack access to personalized and affordable emotional support. This gap leaves many students struggling in isolation, potentially impacting their academic performance, social well-being, and long-term mental health [1]. |
| Key Aspects of Solution | AI-powered journaling, Sentiment Analysis, Mood Tracking, personalised affirmations, and Goal Setting |
| Technology Keywords | Natural Language Processing, Sentiment Analysis, Emotion Detection, Data Encryption, |
| Data Availability | Private |
| Metadata (Type of Data) | Text (journal entries), Emotional labels, Timestamps, Biometric metadata |
| Model Training and Fine-Tuning | Sentiment models leverage pre-trained transformers (RoBERTa [7] and EmpathBERT [8]), with performance assessed through manual validation on carefully curated emotional text samples that simulate real-world diversity in gender, regional language, slang, and socio-economic context. This approach helps identify and mitigate potential bias in early development stages. |
| Testbeds or Pilot Deployments | Under development as part of the StudentSphere platform—internal VIT pilot planned |
| Code Repositories | Not Available |

# 2 Use Case Description:

## 2.1 Description

Student mental issues, such as anxiety, stress and emotional fatigue, increase continuously because academic stressors exist alongside social challenges. The ability to express emotions through journal writing exists, but students usually miss out on interactive guidance and continued monitoring of their progress. Lack of engaging mental well-being solutions that respond to student requirements has emerged as an urgent necessity because mental well-being stands as a key priority. The developers intend to build an AI journaling application that gives students immediate emotional assistance. The application advances static journal writing functions by performing emotional pattern evaluations and providing individualized feedback to users. Through this app, users will gain better emotional management proficiencies to build stronger mental health and emotional strength.. The application leverages NLP techniques to perform multi-label emotion classification, going beyond basic sentiment analysis. It uses fine-tuned transformer-based models such as RoBERTa and EmpathBERT, trained on a custom corpus of student journaling samples reflecting diversity in gender, socio-economic background, regional language, and slang. Instead of relying entirely on external datasets, the models are validated manually using simulated real-world student scenarios to reduce cultural and demographic bias. The system is designed to identify nuanced emotional expressions—including sarcasm, overlapping emotions, and ambiguous phrases—common in student writing. Over time, the AI adapts through continual learning from anonymized user input, generating personalised affirmations, coping strategies, and mood trend insights. Data protection is ensured through end-to-end encryption, and user authentication is safeguarded by biometric verification (facial/fingerprint recognition). All AI decisions are made with transparency, and clinical validation is planned in collaboration with psychologists to ensure ethical deployment and user safety. The solution provides automated support that increases emotional self-awareness and reduces stress by giving users timely assistance. Using this tool gives students the ability to both think about their emotional state and manage their feelings, together with receiving help when they need it. The app provides students with an affordable and private emotional care platform, which helps students focus better academically while fostering their mental wellness.

**Use case status:** Ongoing, in active deployment and continuous development.

**Partners:** *StudentSphere* team at VIT [6]

## 2.2 Benefits of use case

- Promotes emotional regulation, reduces stress and anxiety with AI-based mental health support.
- Enhances students' emotional intelligence and concentration, indirectly boosting learning outcomes.
- Introduces AI innovation in mental health through secure and smart platforms.
- Provides accessible and inclusive emotional tools irrespective of socio-economic background.

## 2.3 Future Works

- Model Enhancements: Expand to detect complex emotions, add voice journaling with speech-to-text, and include multimedia data (emojis, images).
- Collaborations: Partner with EdTech and HealthTech startups, integrate with university student portals, and collaborate with AI mental health research labs.
- Expert Oversight: Involve licensed psychiatrists for AI validation, feedback, and ethical oversight.
- Standardization: Review and standardize emotional markers using psychological scales, Patient Health Questionnaire-9 (PHQ-9) and Generalized Anxiety Disorder 7-item (GAD-7).
- Actions: Offer affirmations, coping tips, or alert wellness teams for high-risk emotions based on inferences.

# 3 Use case requirements

- **REQ-01:** It is critical that sentiment analysis on user-generated journal entries is done.
- **REQ-02:** it is critical that mood pattern detection is done over time.
- **REQ-03:** It is critical that a personalized affirmations are generated using CBT-based guidelines and emotion regulation strategies.
- **REQ-04:** It is expected that User authentication via biometric (face/fingerprint).
- **REQ-05:** It is expected that an end-to-end encryption for stored emotional data will be integrated.
- **REQ-06:** It is of added advantage that a dashboard for mood trends, triggers, and emotional logs.
- **REQ-07:** It is expected that the habit tracking & future message features is integrated.

# 4 Sequence diagram

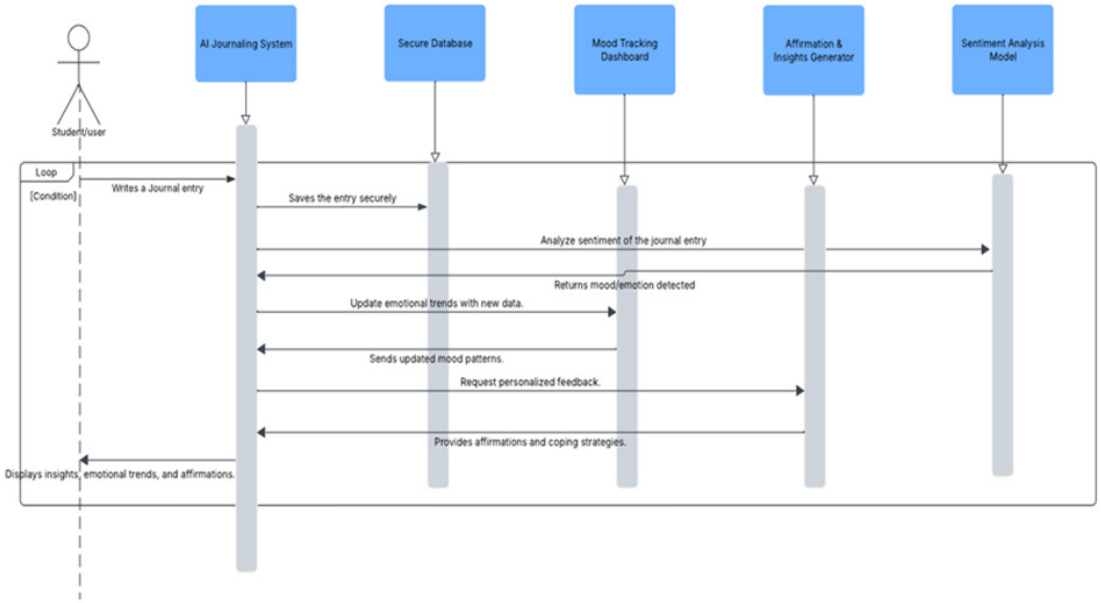



Figure 23.1: Sequence Diagram of our system

The proposed AI-powered journaling system follows the sequence diagram for core operations that support student mental well-being through immediate emotional detection. The AI Journaling System accepts submissions from student users, following which the process continues. The system stores the entry within its secure database to ensure information integrity as well as preserve its confidentiality. The sentiment analysis model identifies core emotional themes from each journal entry, with a target classification accuracy above 85%, emotion-specific precision/recall metrics, and a false-positive rate below 5% for high-risk emotion detection, ensuring timely and reliable feedback within a response time SLA of under 3 seconds. In cases where high-risk emotional signals such as suicidal ideation or self-harm are detected, a crisis detection module is activated. This module triggers a real-time escalation workflow that includes alerting campus counsellors or emergency contacts based on predefined, consent-based protocols and providing users with immediate access to in-app crisis resources such as helplines and support messages. User emotional trends on the Mood Tracking Dashboard get refreshed by the newly obtained sentiment data. The Affirmation & Insights Generator is provided with the new patterns from the platform before initiating an inquiry for targeted feedback. Once they provide information to the generator, it forms personalised affirmations and suitable coping mechanisms in alignment with the emotional state of the users. The user obtains blended synthesised feedback from the loop in the system, which includes emotional trends merged with affirmations and mental health observations for achieving greater emotional awareness and resilience.

## 5 References

[1] Vashisth, A., Kumari, M., & Mishra, A. (2024, September). MoodSync: An AI-Powered Journal for Enhanced Emotional Well-Being. In *2024 International Conference on Artificial Intelligence and Emerging Technology (Global AI Summit)* (pp. 759-764). IEEE.

[2] Mishra, A. K., Bhartiy, K. K., Singh, J., Aluvala, S., Singh, P., & Kishor, K. (2023, September). Unlocking the power of natural language processing through journaling with the assistance. In *2023 3rd International Conference on Innovative Sustainable Computational Technologies (CISCT)* (pp. 1-5). IEEE.

[3] Angenius, M., & Ghajargar, M. (2023, July). Design principles for interactive and reflective journaling with AI. In *Science and Information Conference* (pp. 62-81). Cham: Springer Nature Switzerland.

[4] Angenius, M., & Ghajargar, M. (2022, November). Interactive Journaling with AI: Probing into Words and Language as Interaction Design Materials. In *International Workshop on Chatbot Research and Design* (pp. 150-170). Cham: Springer International Publishing.

[5] Valarmathi, R., Yazhini, G., Narmadha, L. S., & Janani, A. M. (2024, October). Enhancing Mental Well-being through AI-Driven Virtual Environments: The Merrytopia Project. In *2024 International Conference on Power, Energy, Control and Transmission Systems (ICPECTS)* (pp. 1-6). IEEE.

[6] Vellore Institute of Technology, Chennai, India, https://vit.ac.in/campuslife/studentswelfare

[7] RoBERTa, https://huggingface.co/docs/transformers/en/model_doc/roberta

[8] EmpathBERT: A BERT-based Framework for Demographic-aware Empathy Prediction, https://doi.org/10.48550/arXiv.2102.00272

# Use Case 24: Brain-wave driven wheelchair for paralysis patients using EEG and ML

**Organization**: VIT Chennai

**Country**: India

**Contact Person(s)**:

Aaron David Don, aarondaviddon12@gmail.com,

Joshua Banchamin, joshuabanchamin@gmail.com,

Khoushal, khoushal2610@gmail.com,

Pratyaksh Lodhi, pratyakshlodhi2@gmail.com,

Dr. Suganya G, suganya.g@vit.ac.in,

Dr. Nithya Darisini P.S, psnithyadarisini@vit.ac.in.

## 1 Use Case Summary Table

| Item | Details |
|---|---|
| **Category** | **Healthcare** |
| Problem Addressed | Patients with paralysis tend to be too dependent on others for their mobility. Their dependence on a caretaker and their inability to even carry out basic motor motions is a huge strain on their loved ones financially and resource-wise. Moreover, this absolute reliance on others is a huge drain on the mental health of the patients, which in turn deters their chances of healing and living. Our product is a wheelchair which helps paralysis patients move around using just the power of their brain. |
| Key Aspects of Solution | This project harnesses Electroencephalogram (EEG) and machine learning technology to achieve this truly noble system. The EEG is used to collect brainwave data from subjects. This data is pre-processed and used to train an ML model, which classifies a live stream of EEG data into various classes mapped to various actions/motions, which is passed to a microprocessor, which in turn actuates these signals by controlling the relevant actuators. |
| Technology Keywords | EEG, machine learning, signal processing, ESP32, L298, Arduino, Raspberry PI, Internet of Things (IOT). |
| Data Availability | Private |
| Metadata (Type of Data) | EEG-signals: Waves<br>Dataset to train ML model: Text |

(continued)

| Item | Details |
|---|---|
| Model Training and Fine-Tuning | A Support Vector Classifier (SVC) model is trained using the preprocessed and noise-filtered dataset. The model employs a Radial Basis Function (RBF) kernel, which is well-suited for handling the complex patterns present in EEG signals. Initially, a basic SVC model is trained with default parameters. To improve performance, hyperparameter optimization is performed using techniques like GridSearchCV. Parameters such as the regularization term (C) and the gamma parameter are fine-tuned through 5-fold cross-validation. |
| Testbeds or Pilot Deployments | We curated models using data from 10 different people, and testing has shown very promising results. We also have Memoranda of Understanding (MoUs) with hospitals to further test and refine our system. |
| Code repositories | Not Available |

# 2 Use Case Description

## 2.1 Description

Paralysis victims and senior citizens have always been at the mercy of someone else to help them carry out their day-to-day activities in life. This makes their life more miserable, knowing they will never be independent again. We aim to address this issue through the implementation of our brain-controlled wheelchair. Our devices aim to solve all their mobility concerns to make them feel more at ease and regain control of their lives. Our product is a brain wave-controlled wheelchair, which makes use of EEG technology to obtain readings of the brain signals. All of these signals are later preprocessed and then classified using machine learning and fed to the wheelchair, which is then set to perform the various functions as and when they are required. Our product poses an innovative and significantly affordable alternative to the options available in the market for the locomotion of paralysis patients. Upon widespread implementation, it will make the lives of the users much more comfortable. The data is collected from 10 individuals who are made to wear the sensors according to the international 10-20 system. The sensors are placed on the motor cortex to detect any possible movements. Each individual is later asked to behave differently to obtain different kinds of data. At first, they are asked to remain in a calm state – think of only pleasant thoughts and nothing too intense. Later, they are made to maintain a position where they are told to primarily focus on clenching their jaw, and finally, the next set of data is to determine active movements, so they are told to put as much stress as possible into thoughts about limb movement without actually moving them. This is to simulate the conditions of the paralysed people and to train the model based on the respective data.

Each state is maintained for about 20 minutes when the data is continuously collected and then labelled according to what the particular state was. In all, an individual is observed for 60 minutes to obtain 1.8 million data points. When the data from all the participants is combined, the total obtained is 18 million data points. All of these points are labelled according to the user they were collected and the particular state. The data collected is stored as comma-separated value files stored locally on a system. This data is further used in training the model to obtain

proper predictions. Thus, the total training time is an hour, broken down into three 20-minute sessions.

**Use Case Status:** In development Process:

Partners: None

## 2.2 Benefits of Use case

Project outcomes: This wheelchair uses EEG to improve the patient's mobility leading to improved health and quality of life.

The technology mobilizes the paralyzed, helping them have more of a social life by eliminating reliance on others.

The brain-computed wheelchair is a groundbreaking technology, which combines assistive technology and neuroengineering.

## 2.3 Future Work

For future work, we plan on expanding the functionality of the EEG wheelchair to meet more demands. Some of the future implementations for our device:

- Safety features: More safety features, like proximity alert or emergency stop, will be added to the functionalities.
- Integrating EOG: We plan on integrating EOG technology as well to track eye movement as well to help them in many more aspects than just locomotion.
- Linking up with IoT: While our device currently serves its basic purpose to assist people with mobility, upon integration with IoT, we will be able to present a device with unlimited capabilities, from cooking/cleaning/household work to covering all the tasks that need to be done for these people.
- Thermal Throttling: A thermal throttling system can be implemented to ensure that the system adapts to temperature changes. The system can be added using a simple temperature sensor added to the L298N, interfacing it with the Arduino, with a code to monitor the temperature levels. Measures to reduce the heating or to augment system activity can be taken, e.g., a cooling fan, to ensure the smooth operation and durability of the machine.

# 3 Use Case Requirements

- **REQ1**: It is critical that the system accurately acquire raw brainwave (EEG) signals from the user's motor cortex, as per the international 10-20 system, and efficiently preprocesses these signals to remove noise and artefacts for reliable interpretation.
- **REQ2**: It is critical that the system employ a trained machine learning model to classify preprocessed EEG signals into discrete wheelchair commands (e.g., forward, backward, turn left, turn right, stop) with a high degree of accuracy and minimal latency.
- **REQ3:** It is critical that the system translate generated commands into real-time wheelchair movement, ensuring smooth, responsive, and predictable control to enable the user to navigate their environment effectively.
- **REQ4:** It is critical that the system facilitate a structured training process, collecting labeled EEG data from individual users across various mental states (calm, jaw clench,

imagined limb movement) to robustly train and calibrate the machine learning model for personalized control.

- **REQ5**: It is critical that the system efficiently collect, accurately label, and securely store large volumes of EEG data (e.g., 1.8 million data points per user) locally in a structured format, such as comma-separated value (CSV) files, for model training and future analysis.
- **REQ6:** It is critical that the system safety features, including an accessible emergency stop mechanism and real-time monitoring for potential malfunctions or unsafe conditions, to ensure the user's well-being at all times.

## 4 Sequence Diagram

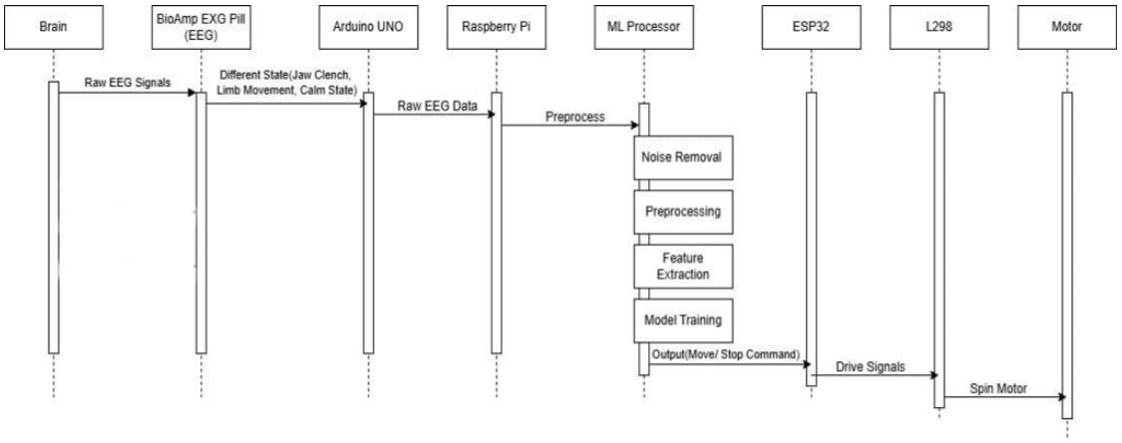



Figure 1: Sequence Diagram of our system

## 5 References

[1.] Hajare, R., & Kadam, S. (2021). Comparative Study Analysis of Practical EEG Sensors in Medical Diagnoses. Global Transitions Proceedings, 2, 467–472

[2.] Zendehdel, N., Chen, H., Song, Y. S., & Leu, M. C. (2024). Implementing Eye Movement Tracking for UAV Navigation. Proceedings of the 2024 International Symposium on Flexible Automation, 2, V001T08A004.

[3.] Basharpoor, S., Heidari, F., & Molavi, P. (2019). EEG coherence in theta, alpha, and beta bands in frontal regions and executive functions. Applied Neuropsychology: Adult.

[4.] Dev, A., Rahman, M. A., & Mamun, N. (2018). Design of an EEG-based Brain Controlled Wheelchair for Quadriplegic Patients. 2018 3rd International Conference for Convergence in Technology (I2CT), 1–5

[5.] Chaddad A, Wu Y, Kateb R, Bouridane A. Electroencephalography Signal Processing: A Comprehensive Review and Analysis of Methods and Techniques. Sensors (Basel). 2023;23(14):6434. Published 2023 Jul 16. doi:10.3390/s23146434

[6.] I. Iturrate, J. Antelis and J. Minguez, "Synchronous EEG brain-actuated wheelchair with automated navigation," 2009 IEEE International Conference on Robotics and Automation, Kobe, Japan, 2009, pp. 2318-2325

[7.] Kumar Chaudhary, A., Gupta, V., Gaurav, K., Kumar Reddy, T., & Behera, L. (2023). EEG Control of a Robotic Wheelchair. IntechOpen. doi: 10.5772/intechopen.110679

[8.] Redwan, S. M., Uddin, M. P., Ulhaq, A., Sharif, M. I., & Krishnamoorthy, G. (2024). Power spectral density-based resting-state EEG classification of first-episode psychosis. Scientific Reports, 14(1), 15154. https://doi.org/10.1038/s41598-024-66110-0

[9.] Manpreet Kaur, Neelam Rup Prakash, Parveen Kalra & Goverdhan Dutt Puri (2019): Electroencephalogram-Based Pain Classification Using Artificial Neural Networks, IETE Journal of Research

[10.] Morley, A., Hill, L., & Kaditis, A. G. (2016). 10-20 system EEG placement. European Respiratory Society.

[11.] Satheesh Kumar, J., & Bhuvaneswari, P. (2012). Analysis of Electroencephalography (EEG) Signals and Its Categorization – A Study. Procedia Engineering, 38, 2525–2536. https://doi.org/10.1016/j.proeng.2012.06.298

[12.] Cruz, A. M., & Wishart, R. (2022). Evaluation of machine learning algorithms for classification of EEG signals. Technologies, 10(4), 79. https://doi.org/10.3390/technologies10040079

[13.] Peng, W. (2019). EEG Preprocessing and Denoising. In L. Hu & Z. Zhang (Eds.), EEG Signal Processing and Feature Extraction. Springer Nature Singapore. https://doi.org/10.1007/978-981-13-9113-2

[14.] Landu Jiang, Cheng Luo, Zexiong Liao, Xuan Li, Qiuxia Chen, Yuan Jin, Kezhong Lu, Dian Zhang, SmartRolling: A human machine interface for wheelchair control using EEG and smart sensing techniques, Information Processing & Management, Volume 60, Issue 3, https://doi.org/10.1016/j.ipm.2022.103262

# Use Case 25: Smart Watch based Integration System for Habituating Face Paralysis Monitoring

**Organization**: Vellore Institute of Technology, Chennai

**Country**: India

**Contact Person(s)**:

Andhavarapu Balu, andhavarapu.balu2022@vitstudent.ac.in

Ishan D. Jain, ishandpalresha@gmail.com

Dr. Geetha S., geetha.s@vit.ac.in

## 1 Use Case Summary Table

| Item | Details |
|---|---|
| Category | Healthcare |
| Problem Addressed | Current methods for detecting facial paralysis often rely on clinical observation, which can lead to delayed diagnosis and intervention. This delay can significantly impact patient outcomes, as early and continuous monitoring of facial movements is crucial for effective treatment and recovery. |
| Key Aspects of Solution | Uses a wrist-worn device with a camera that records the facial mimicry patterns and checks for abnormalities |
| Technology Keywords | Wearable Technology, Remote Monitoring and Alerts, Fog Computing, Cloud, Database, Deep Learning, Computer Vision, IoT |
| Data Availability | public [2] |
| Metadata (Type of Data) | Visual (image and video) |
| Model Training and Fine-Tuning | Computer vision models are to be used for face detection, mimicry identification (facial action unit identification), paralysis detection, and time series analysis. |
| Testbeds or Pilot Deployments | Active, with ongoing deployment and iterative refinement |
| Code repositories | Not Available |

# 2 Use Case Description

## 2.1 Description

Facial paralysis is a condition in which the patient is unable to control the muscles of their face. While some causes, such as Bell's palsy, may resolve spontaneously, others, particularly neurological conditions, require early detection to prevent the progression of symptoms and enable timely medical intervention. Identifying the early signs of facial paralysis can serve as an indicator of neurological disorders such as strokes and multiple sclerosis; infectious conditions such as Bell's palsy, Lyme disease, and Ramsay Hunt syndrome; physical trauma, including injuries from surgery or fractures; the presence of tumours; and other specific disorders such as Guillain-Barré syndrome. [1] We propose the use of a wearable, video-based monitoring system inspired by the design of a smartwatch, which takes advantage of the fact that such a device is naturally directed towards the user's face multiple times throughout the day. Recognising that many patients may not regularly interact with dedicated software, the system is designed to continuously and passively monitor the user's face whenever the watch is turned towards them. Each time the patient raises the watch to check the time, the system automatically captures a brief video of the face. Additionally, the patient is expected to spend approximately one minute per day performing facial mimicry exercises, which allow the AI system to assess facial muscle functionality. To reduce the need for user interaction, which may be a deterrent for some patients, the software is designed to automatically detect whether the patient is engaging in these exercises during the captured footage. This ensures a completely hands-free experience. The software keeps a log of the mimicry exercises performed and their respective timestamps, enabling it to prompt the patient to repeat specific gestures if they have not been performed for some time. The device employs existing mechanisms to detect hand-raising gestures, commonly referred to as the "flip to wake" or "raise to wake" feature. Captured videos are transmitted to a nearby fog-layer computational unit, such as a mobile phone, which subsequently uploads them to cloud servers equipped with specialised deep learning models for facial paralysis detection. These models, developed within the academic research community, are sufficient for analysis as the system does not require any bespoke training per individual patient. However, access to timestamped video history enables the models to compare current facial data with that of previous instances, potentially improving diagnostic accuracy by accounting for the patient's natural facial muscle range and individual baseline.

**Use Case Status:** Existing solutions [2]:

**Partners:** Vellore Institute of Technology, Chennai [7]

## 2.2 Benefits of use case

- Early detection and monitoring of facial paralysis can lead to timely medical interventions, improving patient outcomes and quality of life.
- Developing new technologies, such as AI-based analysis or remote monitoring applications, contributes to medical innovation and accessibility.
- Collaboration between researchers, healthcare providers, and technology developers is essential for implementing effective solutions in facial paralysis diagnosis and treatment.

## 2.3 Future Work

Future works are anticipated to

1) Build AI models for mimicry exercise recognition. There are already existing models for detecting facial paralysis. These models can be combined as per our architecture. To build these models, we need to collect a dataset. A dataset comprising predominantly typical individuals along with a number of samples of individuals with facial paralysis should be adequate for the task of training a model for mimicry recognition.
2) A smart watch with the necessary hardware for collecting the video clips should be built. Existing smart watch technology can be leveraged for building the hand movement recognition system.

# 3 Use Case Requirements

- **REQ-01:** It is critical that the smartwatch be able to collect video of the face
- **REQ-02:** It is critical that the video collected should have depth and/or colour information as required by the AI model
- **REQ-03:** It is critical that the smartwatch have a camera or other such hardware instrument that points at the face to capture the video
- **REQ-04:** It is expected that the smartwatch operate like a regular smartwatch that doesn't have an always-on display, i.e., it should detect and respond to the hand movement pointing it towards their face by turning on the screen
- **REQ-05:** It is critical that a video clip of the face is captured every time the user looks at the watch
- **REQ-06:** It is critical that captured video clips should be classified by AI as a normal face clip or the wearer performing any of the mimicry exercises for face paralysis detection
- **REQ-07:** It is critical that the mimicry exercises performed by the wearer are correctly classified
- **REQ-08:** It is expected that Mimicry exercises are detected even when the user has not explicitly interacted with the watch software to give an indication that they are going to do their exercise
- **REQ-09:** It is critical that the sequence of mimicry exercises be analyzed by the AI for face paralysis
- **REQ-10:** It is expected that the software should keep track of the history of mimicry exercises performed by the user
- **REQ-11:** It is expected that the software prompts the user to do specific mimicry exercises if they have not performed those in a while

## 4 Sequence Diagram

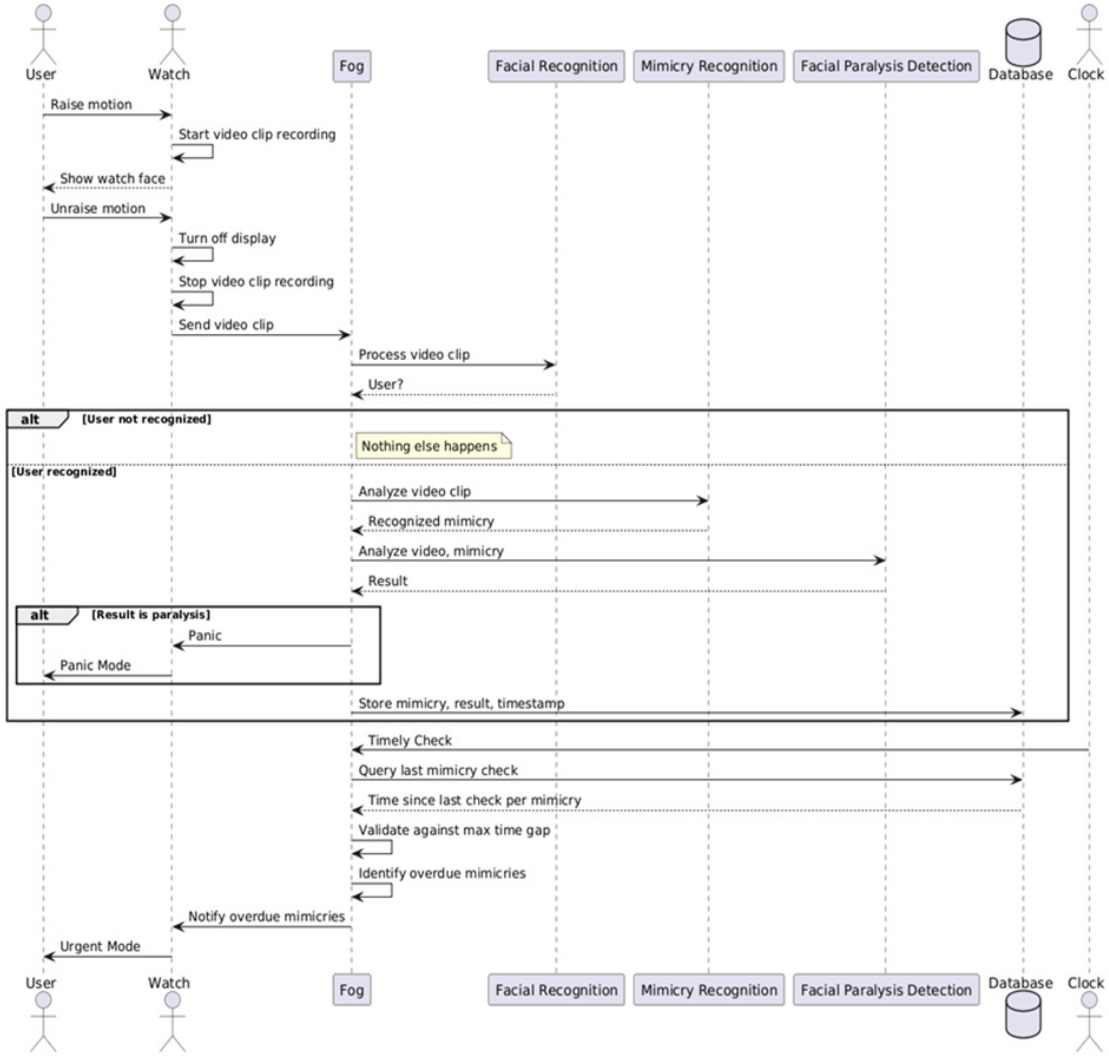



Figure 1: Sequence Diagram of our system

## 5 References

[1] A. Kumar, "Facial Paralysis: Causes, Symptoms, and Treatment," Narayana Health, 2024. https://www.narayanahealth.org/blog/facial-paralysis-causes-symptoms-and-treatment.

[2] M. Mavadati, P. Sanger and M. H. Mahoor, "Extended DISFA dataset: Investigating posed and spontaneous facial expressions," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW)*, Jun. 2016, pp. 1452-1459.

[3] A. et al., Technologies applied in Education and Health, 2021.

[4] R. Guarin et al., "A machine learning approach for automated facial measurements in facial palsy," 2018.

[5] D. Storey et al., "Role for 2D image-generated 3D face models in the rehabilitation of facial palsy," 2017.

[6] D. Watts et al., "Cloud computing mobile application for remote monitoring of Bell's palsy," 2020.

[7] Vellore Institute of Technology, Chennai. https://vit.ac.in/.

# Use Case 26: AI-Enhanced Organoid Platform for Pandemic Preparedness and Personalized Medicine

**Organization:**Cooperative Open Intelligent Computing Industry Alliance (Collaborating with C2itech Limited)

**Country:** China, Hong Kong SAR

**Contact person:**

Alex Li alex.li@c2itech.HK. Phone: +852 98765443 (WhatsApp)

## 1 Use Case Summary Table

| Category | Healthcare |
|---|---|
| Problem addressed | The Artificial Intelligence (AI)-Enhanced Organoid Platform aims to revolutionize pandemic preparedness and personalized medicine by leveraging cutting-edge AI technologies. It integrates 3D organoid models with deep learning algorithms to simulate human respiratory tract responses, predict viral mutations, and optimize drug discovery processes. |
| Solution to the problem | Inefficient Drug Screening: Traditional methods for respiratory virus drug testing are slow and costly, delaying critical treatments during outbreaks.<br>Unpredictable Viral Evolution: Current tools cannot accurately forecast high-risk viral variants, hindering proactive vaccine development.<br>Limited Access to Personalized Medicine: Developing regions lack affordable, tailored treatments for respiratory diseases. |
| Technology keywords | Deep learning models, Generative AI, AI-powered image analysis |
| Metadata (type of data) | a) Viral genome databases, WHO outbreak data; b) C2i's organoid datasets, clinical trial data |
| Data availability | a) Publicly available; b) Privately available |
| Model training | Working on: Pre-training of hybrid BioNeMo + DeepSeekMoE architecture (using ESM-2/Geneformer backbones) on federated biological datasets.<br>Current Focus: Domain-specific fine-tuning for drug discovery workflows (target: **>93% accuracy** by end-Nov). |

(continued)

| Category | Healthcare |
|---|---|
| Testbeds or pilot deployments | 2 biotech partners testing protein function prediction in Hong Kong and Shanghai<br>2 hospital labs validating gene expression analysis in Hong Kong<br>Initial Metrics: Latency <80ms, user satisfaction > 4.7/5 |

# 2 Use Case Summary Table

## 2.1 Description

The AI-Enhanced Organoid Platform aims to revolutionize pandemic preparedness and personalized medicine by leveraging cutting-edge AI technologies. It integrates 3D organoid models with deep learning algorithms to simulate human respiratory tract responses, predict viral mutations, and optimize drug discovery processes.

Problem to Solve:

1) Inefficient Drug Screening: Traditional methods for respiratory virus drug testing are slow and costly, delaying critical treatments during outbreaks.
2) Unpredictable Viral Evolution: Current tools cannot accurately forecast high-risk viral variants, hindering proactive vaccine development.
3) Limited Access to Personalized Medicine: Developing regions lack affordable, tailored treatments for respiratory diseases.

AI Approach:

- Organoid Simulation: Train deep learning models on 3D organoid data to replicate human respiratory tract responses to pathogens like COVID-19 and influenza.
- Viral Mutation Prediction: Use generative AI to analyze viral genome sequences and predict high-risk mutations, enabling preemptive vaccine design.
- Automated Monitoring: Develop AI-powered image analysis tools for non-invasive organoid monitoring, reducing manual labor and improving accuracy.

Benefits:

- Accelerated Drug Discovery: Reduce drug screening timelines by 50%, enabling faster responses to pandemics.
- Proactive Vaccine Design: Achieve 90% accuracy in predicting viral mutations, minimizing outbreak risks.
- Personalized Medicine: Provide low-cost, AI-driven diagnostics and treatment plans for respiratory diseases, especially in underserved regions.

Benefits of the use case:

- The platform directly improves global health by accelerating drug discovery and enabling personalized treatments for respiratory diseases, which are among the leading causes of death worldwide.
- Through open-access AI tools and certification programs, the project trains medical professionals in organoid technology, fostering a new generation of researchers equipped to tackle future pandemics.

- By integrating biotech and AI, the platform drives pharmaceutical innovation, setting new standards for drug development and healthcare delivery.
- Collaborations with WHO, Institut Pasteur, and ASEAN health agencies ensure global knowledge sharing and equitable access to AI-driven healthcare solutions, particularly for developing countries.

### Partners

HKU-Pasteur C2i (Virology expertise); Institut Pasteur (Vaccine R&D); Gavi, the Vaccine Alliance (Global health equity)

### Links

C2itech Website: www.C2itech.hk

C2i Website: www.C2i.hk

## 2.2 Benefits of the use case

- The platform directly improves global health by accelerating drug discovery and enabling personalized treatments for respiratory diseases, which are among the leading causes of death worldwide.
- Through open-access AI tools and certification programs, the project trains medical professionals in organoid technology, fostering a new generation of researchers equipped to tackle future pandemics.
- By integrating biotech and AI, the platform drives pharmaceutical innovation, setting new standards for drug development and healthcare delivery.
- Collaborations with WHO, Institut Pasteur, and ASEAN health agencies ensure global knowledge sharing and equitable access to AI-driven healthcare solutions, particularly for developing countries.

## 2.3 Future Work

- Data collection
- Model development
- Standards development (AI ethics in organoid research)
- Setup reference tools

# 3 Use case requirements

**REQ-01:** Build Federated AI Network: Establish a secure, privacy-preserving AI network to share organoid data across 10+ hospitals in Asia. This will enable collaborative research while complying with international data protection standards, ensuring equitable access to cutting-edge medical insights.

**REQ-02:** Develop Low-Cost Organoid Kits: Create affordable organoid cultivation kits embedded with AI sensors for real-time pathogen detection. These kits will be distributed to underserved regions, democratizing access to advanced diagnostics and personalized medicine.

**REQ-03:** Launch Certification Program: Design and implement an "AI for Organoid Medicine" certification program targeting 1,000 researchers and healthcare professionals by 2026. The program will include online courses, hands-on workshops, and open-access AI tools to foster global expertise in organoid technology.

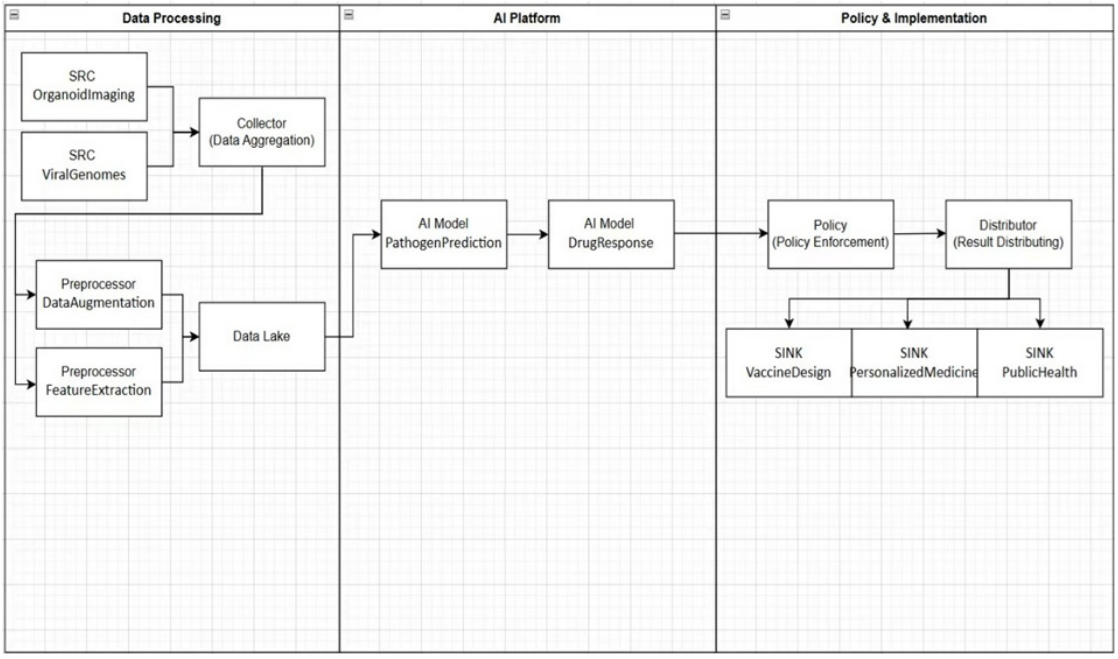
**REQ-04:** Publish Open-Source AI Models: Release AI models for viral mutation prediction and organoid simulation on GitHub, accompanied by detailed documentation and tutorials. This will empower researchers worldwide to adapt and improve the models for local healthcare challenges.

**REQ-05:** Establish Global Standards: Collaborate with WHO and Institut Pasteur to develop ethical guidelines and technical standards for AI-driven organoid research, ensuring responsible innovation and interoperability across institutions.

**REQ-06:** Expand to Other Diseases: Adapt the platform to study non-respiratory diseases (e.g., cancer, neurodegenerative disorders), leveraging the modular design of the AI-organoid system to address diverse medical needs.

**REQ-07:** Engage Public Awareness: Launch a global campaign to educate the public on the benefits of organoid technology and AI in healthcare, fostering trust and encouraging participation in clinical trials.

## 4 Sequence Diagram Design




Claims with Citations:

"AI-based virtual screening reduces drug discovery timelines by 50% by prioritizing high-probability candidates [1-2], significantly cutting costs compared to traditional high-throughput screening [3]."

## 5 References

[1] J. M. Stokes et al., "A Deep Learning Approach to Antibiotic Discovery," Cell, vol. 180, no. 4, pp. 688-702.e13, 2020, doi: 10.1016/j.cell.2020.01.021. Available: https://linkinghub.elsevier.com/retrieve/pii/S0092867420301021 .

[2] A. Zhavoronkov et al., "Deep learning enables rapid identification of potent DDR1 kinase inhibitors," Nat Biotechnol, vol. 37, no. 9, pp. 1038–1040, 2019, doi: 10.1038/s41587-019-0224-x. Available: https://www.nature.com/articles/s41587-019-0224-x .

[3] H. Lin, Y. Wu, J. Chen, S. Huang, and Y. Wang, "(−)-4-O-(4-O- -D-glucopyranosylcaffeoyl) Quinic Acid Inhibits the Function of Myeloid-Derived Suppressor Cells to Enhance the Efficacy of Anti-PD1 against Colon Cancer," Pharm Res, vol. 35, no. 9, p. 183, Jul. 2018, doi: 10.1007/s11095-018-2459-5. Available: https://doi.org/10.1007/s11095-018-2459-5 .

# Use Case 27: Multi-Cancer Early Screening with Non-contrast CT and Deep Learning.

**Organization:** Alibaba Damo (Beijing) Technology Co., Ltd.

**Country**: China

**Contact Person(S)**:

Primary: Yi Chen, elaine.cy@alibaba-inc.com

Secondary: Jianfei Guo, guojianfei.gjf@alibaba-inc.com

## 1 Use Case Summary Table

| Item | Details |
|---|---|
| Category | Artificial Intelligence, Healthcare |
| Problem Addressed | Cancer is a worldwide leading cause of death. Early or incidental cancer detection is associated with prolonged survival and improved patient outcomes. However, early screening of asymptomatic individuals for a wide range of cancers using a single test remains unfeasible due to the low prevalence and potential harms of false positives. Despite the fact that non-contrast computed tomography (CT) offers the potential for large-scale screening, the identification of multi-cancer using non-contrast CT has long been considered impossible, as the low contrast of tumors renders them indistinguishable in non-contrast CT images. Therefore, a deep learning-based early cancer screening solution is introduced to detect and diagnose early multi-cancer from chest and abdominal non-contrast CT, enabling timely treatment with the intent to cure. |
| Key Aspects of Solution | The deep learning-based early cancer screening solution leverages deep learning models to detect and diagnose cancers for eight chest and abdominal organs, including lung, breast, pancreas, stomach, kidney, esophagus, liver and colon, on non-contrast CT scans. Collaborated with clinical experts, we develop deep learning models can help to diagnose multi-cancer with high accuracy and can be readily utilized for opportunistic screening in large-scale asymptomatic patient populations in a fully automatic manner, resulting in safe, low-cost, and effective detection of early-stage malignancies missed by standard-of-care diagnostic techniques and circumventing the toxicity of contrast agent. The deep learning-based early cancer screening solution can help radiologists and medical experts with fast and accurate organ localization, lesion detection and advanced differential diagnosis of lesions, showing great potential for accurate detection of multi-cancer for which no guideline-recommended screening tests are available for average-risk individuals and opening up an exciting possibility of non-invasive, low-cost and universal early multi-cancer detection at both high sensitivity and high specificity levels in which cancers can be detected before symptoms appear. |

(continued)

| Item | Details |
|---|---|
| Category | Artificial Intelligence, Healthcare |
| Technology Keywords | Cancer screening, deep learning, non-contrast CT |
| Data Availability | Private |
| Metadata (Type of Data) | Images, annotations, clinical records |
| Model Training and Fine-Tuning | We combine computer vision models such as 3D U-Net and transformers to segment and classify organs and tumors in CT images, generating both lesion-level detections and patient-level diagnosis. Multi-modal techniques are used to fuse image and clinical records. Hard example mining and incremental learning was used to fine-tune and optimize model for real-world usage, that is, reduce false positives and enable the detection of previously unseen disease types. |
| Testbeds or Pilot Deployments | https://g.alicdn.com/medical-engineering/d3viewer/0.0.174/index.html#/ct2<br>This demo website shows examples of our model detection results on pancreatic, esophageal, gastric, liver and colon cancer. |

# 2 Use Case Description

## 2.1 Description

Pancreatic ductal adenocarcinoma (PDAC), the most deadly solid malignancy, is typically detected late and at an inoperable stage. Early or incidental detection is associated with prolonged survival, but screening asymptomatic individuals for PDAC using a single test remains unfeasible due to the low prevalence and potential harms of false positives. We develop a deep learning approach, pancreatic cancer detection with artificial intelligence (PANDA), that can detect and classify pancreatic lesions with high accuracy via non-contrast CT.

The PANDA model consists of a cascade of three network stages that increase in model complexity and the difficulty level of the tasks performed, involving the first stage for organ localization, the second stage for lesion detection and the third stage for lesion differentiation. The model is integrated with technical innovations including the advanced memory transformer and practical continual learning techniques. PANDA achieved an area under the receiver operating characteristic curve (AUC) of 0.986–0.996 for lesion detection in a multicenter validation involving 6,239 patients across 10 centers, outperforms the mean radiologist performance by 34.1% in sensitivity and 6.3% in specificity for PDAC identification, and achieves a sensitivity of 92.9% and specificity of 99.9% for lesion detection in a real-world multi-scenario validation consisting of 20,530 consecutive patients. DAMO PANDA recently received the FDA's Breakthrough Device designation for early pancreatic cancer detection.

In recent multiple prospective pilot studies, DAMO PANDA screened for over 100,000 individuals and found a total of 16 Pancreatic Ductal Adenocarcinoma (PDAC) that were initially missed by standard-of-care (SOC). Among these, approximately half of them were early-stage PDACs, including one very early-stage tumor (< 1mm in diameter) and one carcinoma in situ.

DAMO PANDA has shown its potential in detecting patients at the early stages of the pancreatic cancer with high precision, enabling timely therapeutic interventions and significantly improving clinical outcomes.

Use Case Status: Pilot

**Category:** Artificial Intelligence, Healthcare

**Partners**: Guangdong Provincial People's Hospital; The First Affiliated Hospital of Zhejiang University School of Medicine; Department of Radiology, Shanghai Institution of Pancreatic Disease; The Affiliated People's Hospital of Ningbo University; Alibaba Foundation

## 2.2 Benefits of use case

1. safe, low-cost, high accuracy, early cancer detection using non-contrast CT scans.
2. detecting cancer before symptoms appear and thereby saving lives.
3. validated using pilot studies and expert feedback, the system enables timely intervention, and significant benefits for prognosis.

## 2.3 Future Work

To achieve efficient and accurate early multi-cancer screening with a single non-contrast CT imaging examination, we propose a deep learning-based early cancer screening solution that is capable of 1) detection of eight major cancers (the top eight with the highest mortality in the world); 2) differential diagnosis of cancers; and 3) measurement and station labelling of eight major organs, in the early-stage with non-contrast CT scans. The proposed deep learning-based early cancer screening solution maximizes the clinical value and patient benefits from a single CT imaging examination, achieving accurate and efficient detection of multi-cancer at both high sensitivity and high specificity levels, which has long been considered impossible due to the low contrast of the tumor presented in CT imaging.

It is possible to build a more accurate AI-based cancer screening model by the integration of routine non-contrast CT imaging with clinical history data in the future. The FDA's recognition of DAMO PANDA signifies a breakthrough in regulatory acceptance of AI technologies, paving the way for their widespread implementation worldwide. We aim to establish collaborations with new partners to validate our model across different regions. More importantly, AI-driven screening should undergo rigorous evaluation equivalent to conventional methods—through randomized controlled trials (RCTs) comparing its efficacy to validated comparators in terms of all-cause mortality reduction. With global partnerships, we wish to achieve the real clinical value of AI-based screening method for cancer in reducing all-cause mortality to address the global cancer burden.

# 3 Use Case Requirements

**REQ-01**: It is required to have routine/non-contrast CT scan (chest or abdomen), DICOM format for radiological examination. To use CT Scanner Manufacturers which comply with DICOM 3.0 protocol.

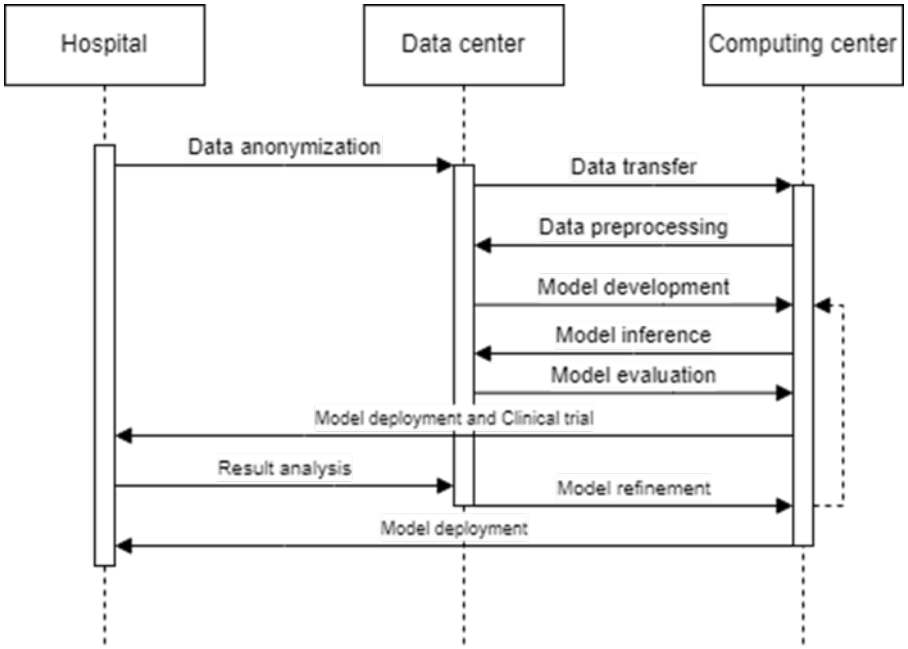
**REQ-02**: It is required to have 16 slices and beyond CT Scanners. CT Scan Parameters: tube voltage 70-140kv, tube current 10-400 mA, slice thickness 1-5mm (> 2mm recommended),

continuous scanning, reconstruction method - chest CT (mediastinal window) or abdominal CT. Resolution 512*512.

**REQ-03**: For hybrid cloud deployment, it is required to have Front-end Machine (CPU): 16 cores, 32GB RAM or higher. Operating System: CentOS 7.9 or a compatible version of Ubuntu 20.04. Public Network Bandwidth: 100 Mbps or higher.

**REQ-04**: For on-premises deployment, it is required to have Inference Machine (GPU): 16 cores, 32GB RAM or higher, Nvidia 3090 or better. Operating System: CentOS 7.9 or a compatible version of Ubuntu 20.04. Public Network Bandwidth: 100 Mbps or higher.

## 4 Sequence Diagram




## 5 References

[1] K. Cao *et al.*, "Large-scale pancreatic cancer detection via non-contrast CT and deep learning," *Nature Medicine*, vol. 29, Nov. 2023, doi: https://doi.org/10.1038/s41591-023-02640-w

[2] J. Yao *et al.*, "Effective Opportunistic Esophageal Cancer Screening Using Noncontrast CT Imaging," *Lecture notes in computer science*, pp. 344–354, Jan. 2022, doi: https://doi.org/10.1007/978-3-031-16437-8_33 .

[3] L. Yao *et al.*, "DeepCRC: Colorectum and Colorectal Cancer Segmentation in CT Scans via Deep Colorectal Coordinate Transform," *Lecture notes in computer science*, pp. 564–573, Jan. 2022, doi: https://doi.org/10.1007/978-3-031-16437-8_54 .

[4] D. Guo *et al.*, "Thoracic Lymph Node Segmentation in CT Imaging via Lymph Node Station Stratification and Size Encoding," *Lecture notes in computer science*, pp. 55–65, Jan. 2022, doi: https://doi.org/10.1007/978-3-031-16443-9_6 .

[5] J. Chen *et al.*, "CancerUniT: Towards a Single Unified Model for Effective Detection, Segmentation, and Diagnosis of Eight Major Cancers Using a Large Collection of CT Scans Cancer Detection in 8 Organs." Accessed: Jun. 09, 2025. [Online]. Available: https://openaccess.thecvf.com/content/ICCV2023/papers/Chen_CancerUniT_Towards_a_Single_Unified_Model_for_Effective_Detection_Segmentation_ICCV_2023_paper.pdf

[6] K. Yan *et al.*, "Liver Tumor Screening and Diagnosis in CT with Pixel-Lesion-Patient Network," *Lecture notes in computer science*, pp. 72–82, Jan. 2023, doi: https://doi.org/10.1007/978-3-031-43904-9_8.

[7] "达医智影," Alicdn.com, 2025. https://g.alicdn.com/medical-engineering/d3viewer/0.0.174/index.html#/ct2 (accessed Jun. 09, 2025

# 4.2. Category 2: Climate Change

## Use Case 1: GeoGPT: AI-driven Marine Ecosystem Monitoring and Protection

**Organization**: Zhejiang Lab

**Country**: China

**Contact Person(S)**:

Primary: Dandan Huang, huangdandan@zhejianglab.org

Secondary: Junhui Shi, junhuishi@zhejianglab.org

Additional Contacts: Ying Zhou, zhouying@zhejianglab.org

### 1 Use Case Summary Table

| Item | Details |
|---|---|
| Category | Climate Change/Natural Disaster |
| Problem Addressed | The use case addresses the challenge of monitoring and protecting marine ecosystems, particularly in areas like deep-sea mineral exploration and harmful marine biological disaster detection, which pose significant environmental risks. |
| Key Aspects of Solution | Use of AI system powered by large language models (LLMs) with retrieval-augmented generation (RAG) technology; AI models for marine ecosystem monitoring and protection; Improved performance in prediction for microalgae distribution and the impacts of climate change on marine ecosystems; Built upon a diversified portfolio—Llama (USA) [3], Qwen (China)[4], Mixtral (Europe)[5], and DeepSeek R1 (China)[6] – GeoGPT offers enhanced multi-modal and scientific discovery capabilities. |
| Technology Keywords | LLMs for marine ecosystem protection; Retrieval-augmented generation (RAG) |
| Data Availability | Public<br>(repository: https://github.com/GeoGPT-Research-Project/GeoGPT) [7] |
| Metadata (Type of Data) | Text; Visual (image recognition for species identification); Environmental data (temperature, water quality, algae distribution) |

(continued)

| Item | Details |
|---|---|
| Category | Climate Change/Natural Disaster |
| Model Training and Fine-Tuning | The GeoGPT models are built upon a diversified portfolio, including Llama (USA), Qwen (China), Mixtral (Europe), and DeepSeek R1 (China), offering users the flexibility to choose the model best suited to their region and needs. These models undergo a series of post-training processes, which include Continual Pre-training (CPT), Supervised Fine-Tuning (SFT), and Direct Preference Optimization (DPO). [3][4][5][6] The training process is carried out in three stages:<br>• Continual Pre-training (CPT): This stage utilizes a diverse set of geoscience-related corpora to obtain a solid specialized model for geoscience.<br>• Supervised Fine-tuning (SFT): This stage enhances the model's ability to follow geoscience-specific instructions by incorporating QA pairs labelled by geoscientists, along with those generated from the training corpus in CPT stage.<br>Human Preference Alignment: This stage uses the Direct Preference Optimization (DPO) with preference data labelled by large language models to align model's responses with human expectations and preferences. |
| Testbeds or Pilot Deployments | GeoGPT platform: https://geogpt.zero2x.org [8]<br>The open platform for GeoGPT, offering intelligent services such as literature analysis, data extraction, map chat, domain-specific Q&A, etc. |

# 2 Use Case Description

## 2.1 Description

GeoGPT is an AI system powered by large language models (LLMs) with retrieval-augmented generation (RAG) technology, applicable to the monitoring and protection of marine ecosystems. The GeoGPT models are built upon a diversified portfolio, including Llama (USA), Qwen (China), Mixtral (Europe), and DeepSeek R1 (China), offering users the flexibility to choose the model best suited to their region and needs. [3][4][5][6]

One of its key applications is in deep-sea mineral exploration. By incorporating mineral prospecting prediction models and risk assessment models, GeoGPT helps define target exploration areas, assess the economic viability of extracting mineral resources, and provide impact analyse to support exploration decision-making. This approach helps minimize unnecessary mining activities, thereby reducing ecological damage. The system currently achieves an accuracy rate of over 60%.

Another major application of GeoGPT is in marine biological disaster monitoring. The system incorporates microalgae species classification algorithms to accurately identify harmful algae, analyze algae distribution patterns, detect red tide signals, and predict the timing, location, and scale of red tide outbreaks, significantly improving monitoring and early warning efficiency. Furthermore, GeoGPT evaluates the impact of global warming on microalgae distribution, shedding light on how climate change influences the behavior, distribution, and interactions of algae in ecosystems. This capability provides critical scientific support for aquatic ecosystem

protection and disaster prevention, particularly in mitigating the impacts of harmful algal blooms, red tides, and other marine biological disasters that threaten marine biodiversity and human activities. GeoGPT currently achieves an accuracy rate of 81.3% in deep-sea species identification and 93.2% in distinguishing biological from non-biological entities. The recognition efficiency is 40 milliseconds per image, a significant improvement over the manual recognition time of 40 seconds per image.

While GeoGPT has shown promising results, it currently faces challenges related to data noise. Ongoing optimizations to model structures and the collection of high-quality data will enable further enhancements in its performance, bolstering its capacity to support marine ecosystem protection and disaster early warning.

This solution evaluates the influence of global warming on marine microalgae distribution and forecasts its impact on aquatic ecosystems. Provide technical support for developing climate-adaptive management strategies, reducing environmental and economic risks associated with climate change. It also reduces unnecessary harm to marine ecosystems by conducting risk assessments in deep-sea mineral resource exploration. Strengthen red tide monitoring and early warning through accurate identification of harmful algae, thereby safeguarding aquatic ecosystem health. Furthermore, it enhances global ecosystem protection and promote sustainable development by fostering technological collaboration and enabling data sharing, improving collective efficiency in ecological preservation efforts. Last, it builds an advanced AI system to deliver innovative solutions for ecological protection and sustainable resource management.

**Use Case Status:** The use case is part of a larger research project and part of a larger product development.

**Partners:** Deep-time Digital Earth (DDE) Big Science Programme; The International Union for Conservation of Nature; Second Institute of Oceanography, Ministry of Natural Resources (MNR); National Deep Sea Center; National Marine Data Center, Department of Nature Resources.

## 2.2 Benefits of the use case

- Evaluate the influence of global warming on marine microalgae distribution and forecast its impact on aquatic ecosystems. Provide technical support for developing climate-adaptive management strategies, reducing environmental and economic risks associated with climate change.
- Reduce unnecessary harm to marine ecosystems by conducting risk assessments in deep-sea mineral resource exploration. Strengthen red tide monitoring and early warning through accurate identification of harmful algae, thereby safeguarding aquatic ecosystem health.
- Enhance global ecosystem protection and promote sustainable development by fostering technological collaboration and enabling data sharing, improving collective efficiency in ecological preservation efforts.
- Build an advanced AI system to deliver innovative solutions for ecological protection and sustainable resource management

## 2.3 Future Work

**Data collection:** Collect high-quality, multimodal geoscience data through open and collaborative methods, ensuring a rich and diverse open-source dataset to support comprehensive model training and improve GeoGPT's effectiveness across various scenarios.

**Model development**: GeoGPT currently operates with parameter scales ranging from 66 billion to 110 billion. Future initiatives aim to further optimize its architecture, create models adaptable to diverse applications and inference devices, and open-source the model series. These efforts will promote collaboration, drive technological innovation, and enhance accessibility to AI-driven solutions.

**Create new variations/extensions to the same use case**: GeoGPT will integrate with other ecological protection tools to provide tailored solutions for diverse environmental challenges. Key improvements include boosting its image recognition and predictive warning capabilities, developing microalgae classification models using microscopic images, and establishing a spatiotemporal prediction system for red tide outbreaks. These advancements will offer scientific support for addressing critical ecological issues.

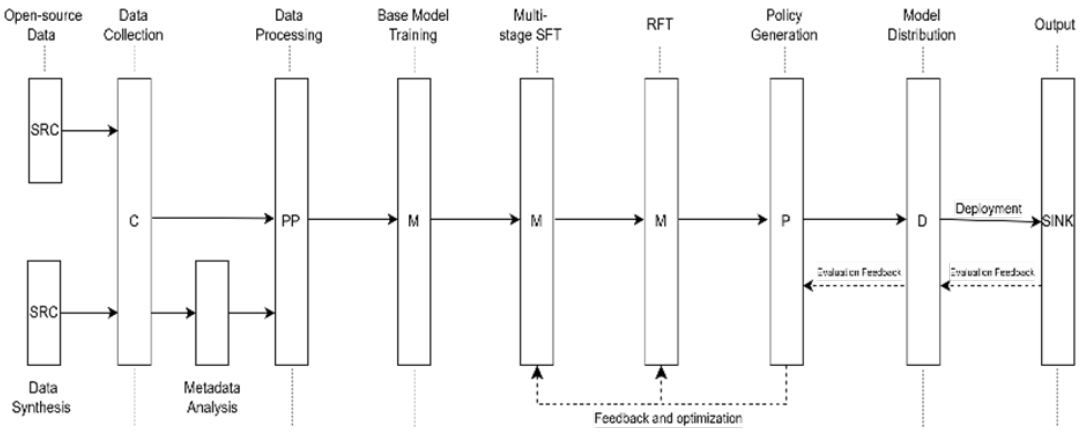
**Standards development related to the use case**: Participate in ITU activities related to Natural Disaster Management, promote the application of AI in the field of marine disaster management, develop standards to provide norms and guidance for the industry, and facilitate the promotion and implementation of best practices.

**Setup reference tools, notebooks and simulation environment:** Provide users with practical reference tools, detailed notebooks, and simulation environments to enhance their ability to utilize GeoGPT for recognition, monitoring, prediction, and decision-making. This will improve the operational efficiency and usability of the platform, empowering users to make data-driven decisions in ecological preservation efforts.

# 3 Use Case Requirements

- **REQ-01**: It is required to have high-quality multimodal geoscience data for model training.
- **REQ-02**: It is required to have scalable cloud infrastructure for AI model deployment.
- **REQ-03**: It is required to have integration with existing marine ecosystem protection tools.
- **REQ-04**: It is required to have real-time data processing and predictive analytics capabilities.

## 4 Sequence Diagram




## 5 References

[1] D. Normile, "Earth scientists plan to meld massive databases into a 'geological Google,'" *Science*, Feb. 2019, doi: https://doi.org/10.1126/science.aax1577 .

[2] "Beyond boundaries: GeoGPT's evolution in geosciences," *AAAS Articles DO Group*, Jul. 2024, doi: https://doi.org/10.1126/science.zz2jbhy .

[3] Meta, "Llama 3.2," *Meta Llama*, 2022. https://www.llama.com/

[4] "Qwen Chat," *Qwen.ai*, 2025. https://chat.qwen.ai/

[5] "Mistral NeMo | Mistral AI," *Mistral.ai*, 2024, doi: https://doi.org/101628137472.ingest.de.sentry.io/4508766640210000 .

[6] DeepSeek, "deepseek-ai/DeepSeek-R1 · Hugging Face," *Huggingface.co*, Jan. 21, 2025. https://huggingface.co/deepseek-ai/DeepSeek-R1

[7] GeoGPT-Research-Project, "GitHub - GeoGPT-Research-Project/GeoGPT: Official release of GeoGPT models.," *GitHub*, 2025. https://github.com/GeoGPT-Research-Project/GeoGPT (accessed Jun. 09, 2025).

[8] "GeoGPT," *Zero2x.org*, 2018. https://geogpt.zero2x.org (accessed Jun. 10, 2025).

# Use Case 2: AI-Powered High-Precision Weather Forecasting Model

**Organization:** Alibaba Damo (Hangzhou) Technology Co., Ltd.

**Country**: China

**Contact Person(s)**:

Yi Chen, elaine.cy@alibaba-inc.com

Ting Wang, wt242148@alibaba-inc.com

## 1 Use Case Summary Table

| Item | Details |
|---|---|
| Category | Climate Change/Natural Disaster |
| Problem Addressed | Although widely-recognized as the most effective weather forecasting methods, Numerical Weather Prediction (NWP) is computationally intensive and has limitations in forecasting timeliness and accuracy over extended forecast periods. Based on numerous weather data, AI based technology is expected to capture complex weather patterns and provide economic and precise weather forecast. |
| Key Aspects of Solution | Established a high-precision regional forecasting model based on a self-developed AI-driven global weather forecasting model to deliver accurate weather forecasts with a 1-kilometer grid resolution on an hourly basis, thereby meeting diverse predictive demands from various industries. [1] |
| Technology Keywords | MAE, Swin Transformer, Neural ordinary differential equations, Multi-modality modeling |
| Data Availability | Private<br>Global datasets primarily depend on the publicly available ERA5 reanalysis data. Regional data are derived from the publicly accessible data of Japan's Himawari satellite. However, regional weather observations and regional reanalysis data are proprietary and cannot be publicly disclosed. |
| Metadata (Type of Data) | Numerical tensors, images |
| Model Training and Fine-Tuning | Utilized a Siamese MAE masked autoencoder based pretraining strategy to provide better initialization and learn robust feature representations hidden within highly volatile weather data. Based on the self-developed global weather model, multi-modal weather data were cooperated in the regional weather model to enable rapid refresh accurate weather forecast. |

(continued)

| Item | Details |
|---|---|
| Category | Climate Change/Natural Disaster |
| Testbeds or Pilot Deployments | Electric Transmission System Operators and Distribution System Operators in China<br>Weather forecasting system for State Grid Shandong Electric Power Co.LTD. |

# 2 Use Case Description

## 2.1 Description

Meteorological forecasting plays a crucial role in addressing climate change by delivering precise weather data that align with industrial needs, particularly in the burgeoning field of renewable energy. However, existing weather forecasts often fall short of meeting the specific requirements of different industries. Traditionally, weather forecasting relies on numerical weather prediction methods, utilizing mathematical and physical equations to simulate atmospheric and oceanic movements. This approach establishes an initial weather state and employs computers to solve the equations for future weather conditions numerically. While effective, this method is computationally intensive, demanding significant processing power. Moreover, it has limits in regional forecasting, including infrequent updates and diminishing accuracy over extended forecast periods.

To address these challenges, this approach harnesses advanced algorithms to transform traditional weather prediction methods. By leveraging AI technology alongside a self-developed global meteorological model, a high-precision regional weather forecasting model has been established. This model integrates diverse local data sources, including regional reanalysis data, real-time meteorological information, radar images, satellite imagery, etc, to enhance the precision and details of forecasts. The result is a system capable of delivering weather forecast with a 1-kilometer grid resolution on an hourly basis, thereby providing specialized predictive support for various industries. [1]

The regional weather forecast model demonstrates significantly superior performance compared to ECMWF-IFS, achieving an average improvement of RMSE (Root Mean Squared Error) of 57.8% on key variables in given region. In terms of the global model alone, it consistently outperforms IFS, achieving an average improvement of 20.5%. In particular, for T2M (air temperature measured at a height of 2 meters above the ground), our model Baguan results in an average enhancement of 26.3%. As the lead time increases, the magnitude of our improvement becomes even more evident. [2]

In order to produce operational weather forecasts, we schedule regular data acquisition services for publicly available data and get proprietary data through ftp services. We store data in the Aliyun OSS system and set up the data-transmission pipeline to the product server for regularly model inference.

**Use Case Status:** Operational

**Partners:** [N/A]

## 2.2 Benefits of use case

In this case, we utilize a Siamese MAE masked autoencoder based pretraining strategy to provide better initialization and learn robust feature representations hidden within highly volatile weather data, enabling precise weather insights. We have also independently developed a foundational weather model with varying resolutions that can generate specialized weather indicators for applications such as power energy, agricultural meteorology, and low-altitude weather forecasting. This contributes to industry, innovation, and infrastructure goals. Additionally, based on our comprehensive weather model, we have developed a reliable model for regional renewable energy forecasting, effectively addressing the long-standing challenge of unpredictability in renewable energy. This contributes to stable grid operation, aids in increasing renewable energy consumption, and further alleviates climate change pressures.

## 2.3 Future Work

Leveraging our latest developed global-regional weather forecasting model, we aim to enhance its functionality and broaden its use in two main areas:

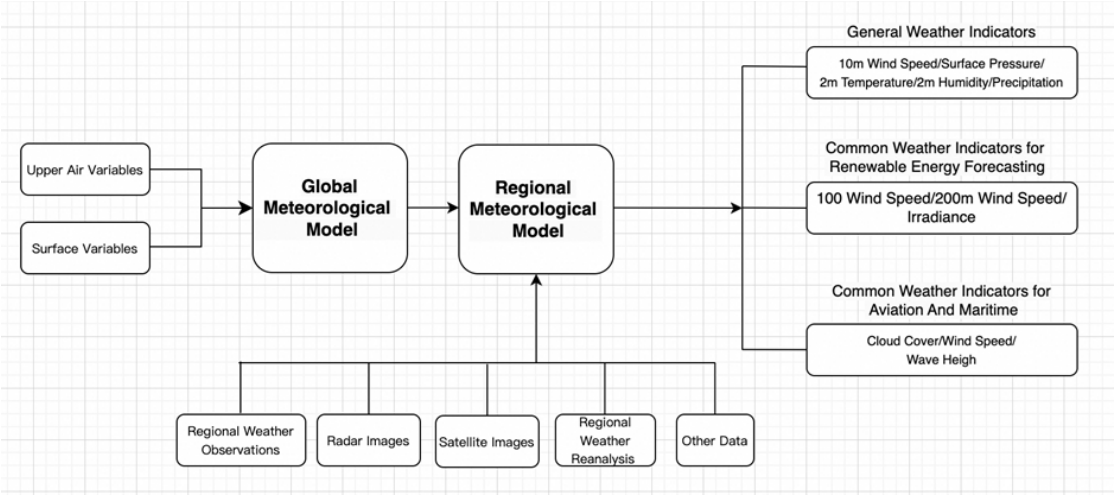
**Technical Development**: We plan to refine our AI-driven weather forecasting model to better address user challenges and societal needs. Our primary focus will be on developing more precise algorithms to predict extreme weather events such as tropical cyclones, heavy rainfall, and snowstorms. This will enhance our ability to forecast the timing and intensity of such events, thereby mitigating their potential societal impact. In light of ongoing climate change, we are also committed to extending our AI-powered weather model to improve forecasting over longer time horizon. This includes building capabilities for climate predictions that cover subseasonal, seasonal, and interannual forecasts. [3]

**Applications**: We aim to promote our AI-enabled global-regional meteorological model both within China and internationally. This will involve offering weather forecasts with finer spatiotemporal resolution to support various industries, including the renewable energy sector and beyond. [4][5]

# 3 Use Case Requirements

- **REQ-01**: It's required to use global weather analysis data for global weather forecasting.
- **REQ-02**: It's required to use regional weather analysis data for regional weather forecasting.
- **REQ-03**: It's required to use regional geography data to enhance model accuracy.
- **REQ-04**: It's recommended to use satellite data and radar image data to enhance forecasting performance.
- **REQ-05**: It's required to use GPU for model training, which is highly recommended for daily online inference.
- **REQ-06**: It's recommended to use cloud-based data transmission for deployment.

## 4 Sequence Diagram




## 5 References

[1] C. Liu, "Alibaba's New AI Model Enhances Weather Forecasting Precision Amid Rising Climate Threats," *Alizila*, Nov. 06, 2024. https://www.alizila.com/alibaba-damo-academy-ai-model-weather-forecasting (accessed Jun. 09, 2025).

[2] P. Niu et *al*., "Utilizing Strategic Pre-training to Reduce Overfitting: Baguan -- A Pre-trained Weather Forecasting Model," *arXiv.org*, 2025. https://arxiv.org/abs/2505.13873 (accessed Jun. 09, 2025).

[3] Y. Guo, T. Zhou, W. Jiang, B. Wu, L. Sun, and R. Jin, "Maximizing the Impact of Deep Learning on Subseasonal-to-Seasonal Climate Forecasting: The Essential Role of Optimization," *arXiv.org*, 2024. https://arxiv.org/abs/2411.16728 (accessed Jun. 09, 2025).

[4] S. Zhang, "5 Ways Alibaba is Leading Sustainability Efforts Demonstrated at COP29," *Alizila*, Nov. 18, 2024. https://www.alizila.com/5-ways-alibaba-is-leading-sustainability-efforts-demonstrated-at-cop29/

[5] "MindOpt_Decision Intelligence – Alibaba Cloud," *Aliyun.com*, 2025. https://opt.aliyun.com/weather (accessed Jun. 09, 2025).

# Use Case 3: AI-powered Monitoring of Mangrove Biomass and Carbon Stocks

**Organization**: data354

**Country**: Côte d'Ivoire

**Contact Person(s)**:

**Primary Contact**: Fabrice ZAPFACK – fabrice.zapfack@data354.co

**Secondary Contact**: Gabriel FONLLADOSA – gabriel.fonlladosa@data354.co

**Additional Contacts**: Salomon KOUASSI – salomon.kouassi@data354.co, Therence TEMFACK – therence.temfack@data354.co

## 1 Use Case Summary Table

| Item | Details |
|---|---|
| Category | Climate Change/Natural Disaster |
| Problem Addressed | Rapid degradation of mangroves due to climate change and human activity; lack of scalable, accurate monitoring tools for biomass and carbon stocks limits effective conservation and policy action. |
| Key Aspects of Solution | – Use of Sentinel-1 (radar), Sentinel-2 (optical), and GEDI LiDAR<br>– AI models (Random Forest, XGBoost, CNNs) for biomass estimation<br>– Calibrated with 300 field inventory plots<br>– Near-real-time monitoring system for conservation and policy use |
| Technology Keywords | AI for Remote Sensing, Multisensor Fusion, Biomass Estimation, Carbon Monitoring, LiDAR, Optical Imagery, Radar, Forest Conservation |
| Data Availability | – Field data: Private (to be made public by May 2025)<br>– Satellite data: Public (Sentinel-1/2, GEDI) |
| Metadata (Type of Data) | – Qualitative: Species name<br>– Quantitative: DBH, tree height<br>– Geospatial: GPS coordinates<br>– Satellite imagery: optical, radar, LiDAR |
| Model Training and Fine-Tuning | – Supervised ML with field-validated features<br>– Random Forest, XGBoost, with potential CNN-based models for improved feature extraction |
| Testbeds or Pilot Deployments | – Côte d'Ivoire (Fresco and Sassandra mangroves)<br>– Senegal (new pilot for tropical forests, 2025–2026) |
| Code Repositories | To be released by end of May 2025 on GitHub |

# 2 Use Case Description

## 2.1 Description

Mangroves play a crucial role in storing carbon—up to four times more than terrestrial forests—and protecting coastal areas. Côte d'Ivoire has lost 95% of its mangroves in just 50 years. Traditional monitoring approaches relying on manual surveys and optical imagery face serious limitations due to cloud cover and cost.

This project introduces an AI-powered system combining:

- Radar (Sentinel-1), optical (Sentinel-2), and LiDAR (GEDI) satellite imagery,
- Supervised machine learning models trained with 300 ground-truth inventory plots,
- Scalable, near-real-time data outputs including biomass maps and carbon estimates.

Output support:

- Conservation actors and government agencies,
- Integration of mangrove data into carbon markets (e.g., REDD+),
- Climate resilience by helping vulnerable coastal communities respond to mangrove loss.

**Use Case Status**: Ongoing pilots in Côte d'Ivoire (2024–2025), expansion to Senegal underway (2025–2026)

Partner

- data354 – Lead partner (AI model development, field integration, remote sensing)
- Lacuna Fund – Funding via "Climate and Forests Machine Learning Datasets" program
- ActumDev – Local field data collection teams in Côte d'Ivoire

## 2.2 Benefits of Use case

- Scalable monitoring to support climate mitigation via accurate carbon stock estimates and early detection of deforestation.
- Supports integration of mangrove conservation into verified carbon markets and builds climate resilience through informed planning.
- Facilitates timely, data-driven decisions to restore mangrove ecosystems critical for coastal protection and emission reduction.

## 2.3 Future Work

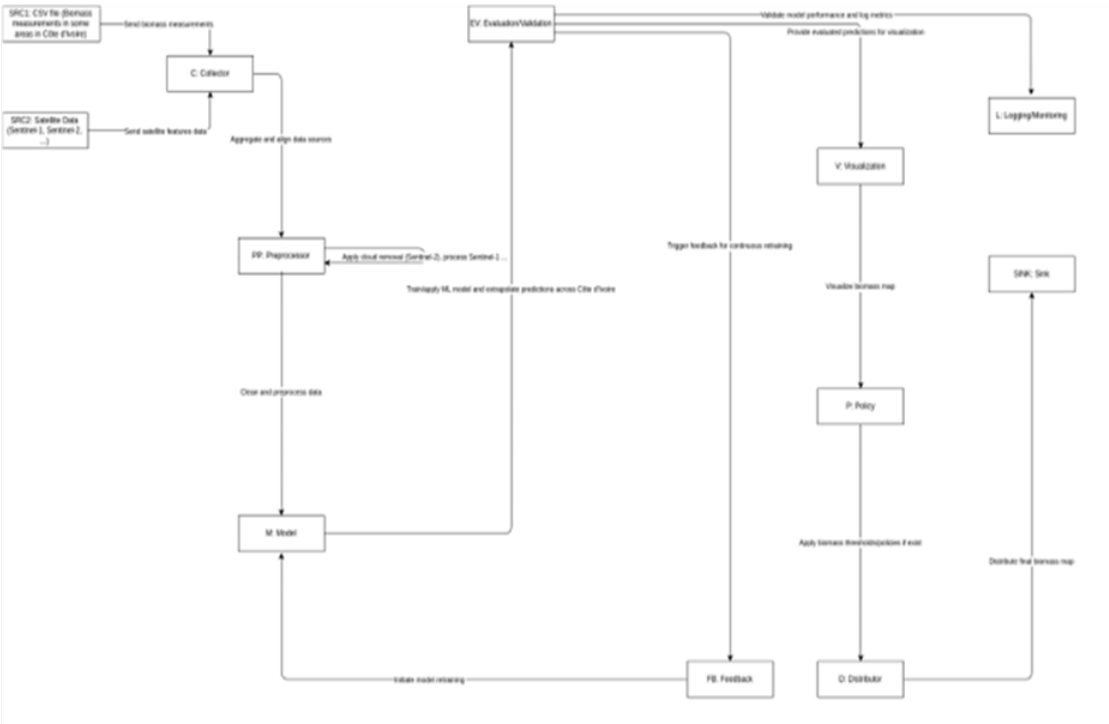
- Certification Alignment: Align biomass estimations with Verra and Gold Standard protocols for carbon credit eligibility
- Geographic Expansion: Rollout to Senegal's tropical forests and broader West Africa
- Model Enhancement: Use of additional data and deep learning (e.g., CNNs) for improved accuracy
- Early Warning Systems: Real-time alerts for degradation events

**Capacity Building**: Collaborations with governments and universities for knowledge transfer and local expertise development

## 3 Use Case Requirements

- **REQ-01**: Access to high-resolution multispectral, radar, and LiDAR satellite data (Sentinel-1, Sentinel-2, GEDI)
- **REQ-02**: Standardized field data collection protocols
- **REQ-03**: Scalable cloud-based infrastructure for AI model training and geospatial computation
- **REQ-04**: Open data platforms for publishing datasets and models (e.g., Kaggle, Hugging Face)
- **REQ-05**: Multi-stakeholder coordination with governments, NGOs, and universities

## 4 Sequence Diagram




## 5 References

[1] Kauffman, J.B. & Donato, D.C. (2012). CIFOR Working Paper 86: Biomass and Carbon Stocks in Mangrove Forests

[2] Copernicus Open Access Hub – Sentinel-1 & Sentinel-2 Satellite Data

[3] NASA GEDI Mission – Global Ecosystem Dynamics Investigation Documentation

[4] Langner et al. (2024). *Nature Scientific Data, 11, 147*. https://doi.org/10.1038/s41597-024-03930-9

[5] Data354 (2024). Pioneering Africa's Biomass Estimation with AI: Scaling Remote Sensing and Machine Learning for Forest and Mangrove Conservation

# Use Case 4: AIrth – AI for Energy Efficiency and Emissions Reduction

**Organization**: Airth

**Country**: Georgia

**Contact Person**:

Lado Bitsadze – bitsadzelado@gmail.com

## 1 Use Case Summary Table

| Item | Details |
|---|---|
| Category | Climate Change/Natural Disaster |
| Problem Addressed | Lack of accessible, real-time tools to reduce energy waste and carbon emissions in homes and businesses. |
| Key Aspects of Solution | - Real-time energy monitoring<br>- AI-based forecasting and anomaly detection<br>- Personalized energy-saving recommendations<br>- Integrated carbon credit marketplace |
| Technology Keywords | Artificial Intelligence, Machine Learning, Energy Optimization, IoT, Carbon Footprint Analysis |
| Data Availability | Private: Energy usage data via smart meters or IoT sensors |
| Metadata (Type of Data) | Time-series data, sensor readings, user interaction logs |
| Model Training and Fine-Tuning | Supervised learning and time-series forecasting (e.g., LSTM); model adapts based on user behavior and seasonal patterns |
| Testbeds or Pilot Deployments | Prototype tested with simulated data; pilots in planning with SMEs in Georgia |
| Code Repositories | GitHub: AIrth |

# 2 Use Case Description

## 2.1 Description

**AIrth** is an AI-powered platform that enables households and businesses to improve energy efficiency and reduce emissions. In Georgia and many parts of the world, energy inefficiency persists due to the lack of real-time, intelligent consumption insights.

AIrth tackles this by integrating smart meter and IoT sensor data to:

- Monitor energy consumption in real time
- Forecast future usage with time-series AI models (e.g., LSTM)
- Offer smart, personalized energy-saving tips
- Support emission tracking and link users to carbon marketplaces

This tool empowers users to reduce costs while actively contributing to global climate goals. It also equips city planners and sustainability officers with dashboards and datasets for policy insights and urban energy strategy planning.

**Use Case Status**: Prototype stage; preparing for pilot deployments with SMEs and residential users in Georgia

### Partners

**Google** – Initial support during the "Build with AI Hackathon" (cloud infrastructure, mentorship)

**Planned** – Local SMEs, municipalities, sustainability-focused NGOs, utility providers, and academic institutions for pilot and scaling

## 2.2 Benefits of use case

- Improves energy efficiency and democratizes access to smart energy insights
- Leverages AI and IoT to drive sustainable innovation
- Helps reduce emissions in urban settings and supports climate-smart planning
- Encourages responsible energy use through data-driven recommendations
- Enables emission reduction and carbon footprint awareness through real-time tracking and insights

## 2.3 Future Work

AIrth will undergo staged expansion, including:

- Pilot Programs with SMEs and residential users to validate performance and behavior impact
- Real-world Data Collection to fine-tune prediction models and refine user feedback mechanisms
- Modular Dashboards for policy makers and energy managers
- Smart Appliance Integration to automate response to AI recommendations

- Open-Source Toolkits and public documentation to encourage adoption in other emerging markets

**Funding and Scaling** through partnerships with investors, utilities, and regional governments

## 3 Use Case Requirements

- **REQ-01**: Integration with smart meters or IoT energy monitoring devices
- **REQ-02**: Scalable and secure cloud infrastructure (e.g., Google Cloud)
- **REQ-03**: Intuitive user interface for mobile/web platforms
- **REQ-04**: Compliance with data privacy and security standards
- **REQ-05**: Georgian language and grid localization support

## 4 Sequence Diagram

Sequence Diagram link

## 5 References

[1] International Energy Agency (2021). *Energy Efficiency 2021*. IEA Report

[2] IEA (2021). *Global Status Report for Buildings and Construction*. IEA Report

[3] ITU-T Y.3172 (2019). *Architectural Framework for Machine Learning in Future Networks*. ITU Recommendation

# Use Case 5: LLM-driven creation of natural hazard geodatabase from digital mass media

**Organization**: HSE University

**Division**: Faculty of Geography and Geoinformation Technology

**Country**: Russian Federation

**Contact Person(s)**:

Anna Derkacheva (aderkacheva@hse.ru)

Maria Sakirkina

Gleb Kraev

Tatiana Aniskina

## 1 Use Case Summary Table

| Item | Details |
|---|---|
| **Category** | Climate Change/Natural Disaster |
| **Problem Addressed** | Traditional natural hazard (NH) databases are fragmented, manually maintained, and hardly accessible, limiting real-time awareness and disaster planning. |
| **Key Aspects of Solution** | – LLM-based automated information extraction from digital news<br>– Geocoding and GIS integration<br>– Capture of localized events<br>– QA and verification pipeline |
| **Technology Keywords** | LLM, Natural Hazard Monitoring, Geodatabase, Text Mining, GIS, Disaster Risk Management |
| **Data Availability** | Yes – demo datasets available online (in Russian) |
| **Metadata (Type of Data)** | Natural hazard type, location, date/time, socio-economic impact, emergency response, event intensity |
| **Model Training and Fine-Tuning** | Custom prompts for LLMs applied to Russian-language news; additional QA filters for validation |
| **Testbeds or Pilot Deployments** | Phase 1 processed 8 million articles → 52,000 NH-relevant texts → 38,000 events (Russia, 2018–2024) |
| **Repository URLs** | Demo Dataset & Project Background [1] (in Russian) |

# 2 Use Case Description

## 2.1 Description

Creating a comprehensive view of natural hazard (NH) exposure requires standardized and frequently updated databases. In Russia and many countries, such data is fragmented, siloed, and manually updated—often unavailable to researchers, planners, or emergency responders. This use case leverages a generative Large Language Model (LLM) to extract structured data from unstructured digital mass media sources. The system processes millions of articles in Russian, identifies relevant natural hazard reports, geocodes them, and transforms them into structured geospatial entries for GIS and disaster risk frameworks.

Approach Highlights

LLM-powered text mining from over 8 million digital publications

Extraction of hazard type, date, location, and socio-economic impact

GIS-compatible event mapping

Built-in quality scoring and transparency for each event

This enables national-to-local event coverage and supports disaster risk reduction, urban planning, and emergency preparedness initiatives.

Use Case Status: Operational pipeline with an actively updated dataset; part of a larger research and development effort at HSE University.

Partners

- No external partner organizations reported at this time. All development and deployment are led by the internal team at HSE University.

## 2.2 Benefits of use case

- Strengthen resilience and adaptive capacity to climate-related hazards and natural disasters.
- Reduce the number of people affected and the economic losses caused by disasters.

## 2.3 Future Work

Archival data will be expanded by mining earlier news records. The team plans to enhance QA filtering, LLM prompts accuracy, and geocoding precision. Transfer of the resulting database and derived analytics to the interested parties is assumed.

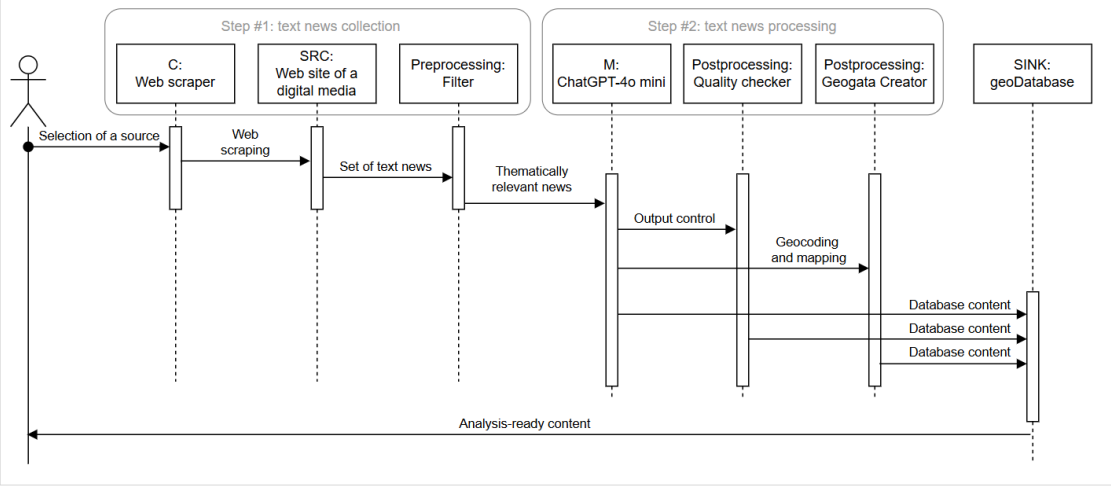
# 3 Use Case Requirements

- REQ-01: A selected Internet source of news publications should not have the legal and technical restrictions for the automatic text scraping.
- REQ-02: After a set of texts is collected in any way, the Internet connection is required for the usage of the designed processing pipeline as is. Currently, two operations are based on the external online services: the language model ChatGPT, and the geocoding

platform Nominatim. Both can be replaced by the offline solutions, if the computing power is sufficient.

- REQ-03: If a proprietary language model is used, it usually requires registration and fee payment proportionally to the volume of data processed. Despite the proposed ChatGPT-4o-mini is very cheap in usage, a project budget should be envisaged.

## 4 Sequence Diagram




## 5 References

[1] HSE University. *Geohazard Project Demo Page*. In Russian. Accessed August 25, 2025. https://geography.hse.ru/georisks/results.

[2] United Nations. "Goal 11: Sustainable Cities and Communities." Accessed August 25, 2025. https://sdgs.un.org/goals.

[3] United Nations. "Goal 13: Climate Action." Accessed August 25, 2025. https://sdgs.un.org/goals.

# Use Case 6: AI-powered Precision Recycling System (WIMPLE)

**Organization**: Seoreu. Co., Ltd.

**Country**: SOUTH KOREA

**Contact Person(s)**:

Jonghyuck LEE, jhlee@seoreu.com

Yeoun Joo LEE, yjlee@seoreu.com

## 1 Use Case Summary Table

| Item | Details |
|---|---|
| Category | Climate Change/Natural Disaster |
| Problem Addressed | Once recyclable waste is mixed with general waste, it becomes extremely difficult to recycle it due to the significant costs involved in separating recyclable materials from mixed waste. To improve these rates, sorting at the source is essential and must be done with high accuracy and precision.<br>The main reason for poor sorting practices lies in the complexity of recycling guidelines, which vary by country and even by region–and frequently change. High-precision sorting, essential for resource circulation-the ultimate goal of recycling, requires active participation from all citizens, which is virtually impossible. As a result, high-cost systems like deposit-refund schemes are being adopted to encourage participation and increase recycling rates. However, it not only covers part of the recycling waste but also costs too much to maintain the system. |
| Key Aspects of Solution | Most recyclable materials generated in daily life are disposable containers. To achieve accurate sorting, it is first necessary to classify these containers by material type–paper, plastic, metal, glass, etc. The purer the material grouping, the higher its value as a recyclable resource. Further classification is also needed, e.g., aluminum vs. steel cans, cardboard vs. carton paper. Additionally, it must be determined whether each item is recyclable and whether additional processing is required.<br>AI enables this granular classification and communicates the sorting information to various robotic systems to ensure precise and automated separation. The system can adapt to different national standards and be applied to deposit-refund schemes. |
| Technology Keywords | Object detection, feature matching, object recognition |
| Data Availability | Public (developed by Seoreu Co, Ltd) dataset: https://www.aihub.or.kr/aihubdata/data/view.do?currMenu=115&topMenu=100&aihubDataSe=data&dataSetSn=71385 [4] |

(continued)

| Item | Details |
|---|---|
| Metadata (Type of Data) | Visual (spatial data) |
| Model Training and Fine-Tuning | The AI model is initially pretrained on the MS COCO dataset to establish general object detection capabilities. It is then fine-tuned using domain-specific data collected from real-world waste materials, including cans, beer cans, plastics, and paper. For regional adaptation, such as in Canada, fine-tuning further incorporates localized data to enable fine-grained differentiation between non-alcohol cans, alcohol cans, US cans, and bi-metal cans.<br>An operational feedback loop continuously gathers new field data to support periodic model updates and performance improvements. |
| Testbeds or Pilot Deployments | https://www.bisco.or.kr/undershop/03_jag/jag01.asp [5]<br>https://www.sedaily.com/NewsView/29X575U740 [6] |

# 2 Use Case Description

## 2.1 Description

Plastics are produced from crude oil, and paper is made from pulp. However, it is not sustainable to continuously consume crude oil and pulp for the use of plastics and paper in our daily lives. Therefore, we must realize circularity, where plastics are recycled into plastics and paper into paper.

To achieve true resource circularity, a sufficient amount of high-quality raw materials must be collected. For this, it is necessary to enable easy recycling (STEP I) and precise sorting (STEP II). Waste must be separated into a form that is as close as possible to its pure raw material state to be recyclable. If precise sorting becomes possible, a sustainable circular economy can be realized.

In one region of Japan, citizens sort waste into as many as 45 different categories, achieving a recycling rate of over 80%. However, it is unrealistic to expect every citizen to follow such a complex sorting system. As a result, most regions simplify their regulations, but compliance rates still remain low. Deposit-refund schemes have been introduced to improve collection rates for recyclables, but the high processing costs due to manual labor remain a significant problem.

Our goal is not only to increase the recycling rate but also to secure high-purity recyclable materials and to realize true resource circularity. Through artificial intelligence (AI), we aim to achieve this by enabling accurate, rapid, uninterrupted sorting and dramatically reducing costs through automation. WIMPLE, installed at Busan Facilities Corporation, has made recycling collection possible in public spaces where it was previously considered impossible. (STEP I)

Waste generated in public places was not only difficult to separate manually afterward but was also often collected already contaminated, rendering separation meaningless. By using WIMPLE to separate recyclable waste from general waste directly at the point of disposal, recycling has been made significantly easier, and the issues of contamination and decline in recycling value have been substantially improved. Moreover, the data collected from the device

has been utilized to calculate the amount of carbon reduction achieved, thereby raising public awareness about the value of resource circulation. It has also enabled optimization of collection schedules for those responsible for waste collection. This represents an innovative case that demonstrates the potential of AI to realize resource circularity in public spaces, where it was previously considered virtually impossible.

Once recyclables are separated from general waste, the next step is to sort them precisely according to their intended use. For this, classification according to national or regional recycling regulations is required. However, due to differences in regulations across countries, localized adaptation of AI models is necessary. We developed the WEE (Waste Evaluation Engine) algorithm package to enable rapid and cost-effective application of local recycling standards. By identifying the needs of the Canadian container deposit refund market – which had not been able to adopt AI due to complex regulations – we applied WEE and succeeded in building a container-sorting AI system achieving over 98% accuracy within just three months. (STEP II)

Lastly, beyond fine classification, we recognize the need for precision separation by material composition. Currently, WEE is capable of most fine-grained classifications. However, we are developing an advanced plastic composition classification system to overcome the final challenge of identifying material compositions at a granular level, with the goal of launching a commercial system within the next year.

WEE has been developed to flexibly adapt to regulations by country or intended use. Our system is designed to quickly and accurately respond to any regulatory framework worldwide. We believe that if our system is widely adopted globally, achieving a recycling rate of 60% in the short term and over 80% in the long term is fully possible- along with improvements in recycled material quality and sustainability.

This solution prevents municipal waste from entering the environment through efficient treatment and recycling, thereby reducing water and soil pollution and protecting ecosystems, improves urban waste management and air quality, reduce GHG emissions from transport and treatment, and enhance infrastructure sustainability and eco-friendly urban development, minimizes waste generation and promote reuse through well-separated recyclables. Optimize plastic waste sorting to increase recycled plastic usage and support circular economy goals, and reduces land-based marine pollution by using automated sorting systems to keep plastics out of oceans, preserving marine ecosystems and enabling sustainable ocean use. Higher recycling rates mean less landfill and incineration, reduced transport emissions, and more circular resource flow. This benefits air, soil, and marine environments. The Busan Facilities Corporation project proved that recycling is possible in previously unsorted areas, resulting in higher recycling rates and better-quality secondary materials.

Use Case Status: Pilot

**Partners:**

Partner I: Busan Facilities Corporation (https://www.bisco.or.kr/)

Busan Facilities Corporation identifies and validates the value generated by our system. Installed at Busan's busiest location, the organization worked with us to resolve real-world challenges

and stabilize operations. Even with many unfamiliar users, citizens became proficient in using the system after about one year.

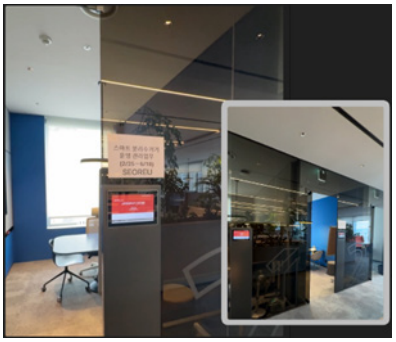
Similar trials have begun in multiple locations such as Busan Central Bus Terminal and Busan Eco-Delta Smart Village.

Recently, a display module was added to transform unsightly waste areas into engaging educational zones where people can experience AI technology and feel pride in their environmental contribution. This reflects o ur joint goal of creating new social and environmental value.

Partner II: Shinsegae Property

Similar projects have begun with Hyundai Department Store Group, one of Korea's top three retail chains.

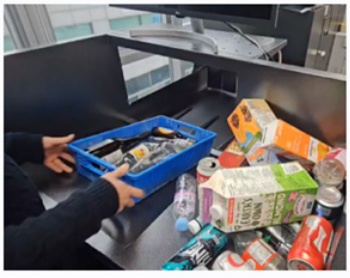
Seoreu Office in Shinsegae

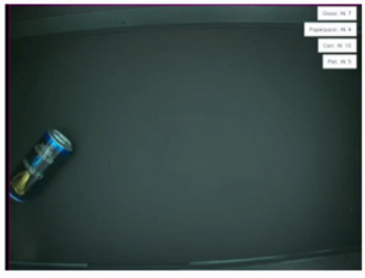


Installation

Partner III: WOORI (Canada)

British Columbia's Return-It program operates over 180 collection centers, recovering 1.8 billion containers annually. Despite the scale, sorting remains manual due to complex classification rules. Our Canadian partner supports local data collection, promotes our solution, and manages local deployment and maintenance.





- ≥ 100 items per minutes
- Operation 24/7
- ≥ 98% accuracy

AI Engines for the WEE

The WEE serves as the core AI platform for precision recycling operations.

It consists of multiple specialized modules designed to achieve real-time high-accuracy waste classification and ensure compliance with local recycling regulations.

Model Structure and Components

WEE is composed of the following key modules:

- Object Detection Module: Detects and localizes visible waste objects from images captured at the recycling stations.
- Feature Matching Module: Extracts and analyzes fine-grained visual features such as texture, color, shape, and labeling patterns to differentiate between different material types.
- Fine-grained Classification Module: Maps the detected objects to specific recycling categories according to region-specific regulations, based on subtle visual differences.
- Object Tracking and Counting Modules: Tracks detected objects over time and counts items for operational reporting. These modules use pretrained models optimized for visual tracking tasks and are not subject to retraining.
- Adaptive Feedback Loop: Collects operational data (e.g., misclassifications, novel material types) and uses it to refine detection and classification modules through periodic retraining.

Pipelines for Training, Deployment, and Update

- Training:

  WEE is trained through supervised learning on a large, annotated dataset of real-world waste images, capturing various material types, contamination levels, and region-specific labeling variations.

- Deployment:

  After training, models are containerized and deployed on edge devices equipped with GPU acceleration for high-speed inference directly at recycling sites.

- Update:

  New data collected from field operations – especially misclassified items or newly encountered materials – is aggregated, annotated, and incorporated into retraining cycles. This ensures that WEE continuously adapts to operational conditions without disrupting service.

Model Verification in Relation to Local Regulations

Before deployment, WEE models are rigorously verified against the recycling rules and regulations specific to the target region:

- A reference dataset reflecting local regulatory standards is curated.
- Model outputs are validated by comparing classification results against the reference set.
- Pilot deployments are conducted, during which model predictions are manually audited by local waste management authorities or regulatory partners.
- This ensures that WEE consistently aligns with local recycling standards and can accommodate regulatory updates over time.

Fine-tuning for Country-Specific Requirements

To meet the needs of diverse regulatory environments, WEE is customized and fine-tuned for each country:

- Label Mapping Adjustment:

The classification schema is modified to reflect material categories defined by local recycling rules.

- Feature Matching Enhancement:

  In Canada, for example, WEE was specifically enhanced to differentiate between:

  - Non-alcohol cans (e.g., soda cans)
  - Alcohol cans (e.g., beer cans)
  - US cans (imported beverage cans subject to distinct deposit regulations)
  - Bi-metal cans (composite cans combining steel and aluminum)

  Fine-grained feature recognition, focusing on subtle differences in material surface, texture, and labeling, was integrated to ensure high classification accuracy.

- Rule-based Constraints:

  Region-specific rules are also incorporated into WEE through a rule-based logic layer, enforcing hard constraints where necessary.

Operational Feedback and Continuous Improvement

Feedback from pilot deployments, audits, and field operators is systematically collected and used to refine WEE's performance:

- Operational feedback identifies new waste types, emerging materials, and classification challenges.
- Audit results are analyzed to adjust detection thresholds or feature matching parameters.
- Regulatory feedback is used to update classification rules as recycling standards evolve.

Through this dynamic feedback loop, WEE maintains strong alignment with country-specific regulations while continuously improving operational robustness.

## 2.2 Benefits of use case

- Prevent municipal waste from entering the environment through efficient treatment and recycling, thereby reducing water and soil pollution and protecting ecosystems.
- Improve urban waste management and air quality, reduce GHG emissions from transport and treatment, and enhance infrastructure sustainability and eco-friendly urban development.
- Minimize waste generation and promote reuse through well-separated recyclables. Optimize plastic waste sorting to increase recycled plastic usage and support circular economy goals.
- Reduce land-based marine pollution by using automated sorting systems to keep plastics out of oceans, preserving marine ecosystems and enabling sustainable ocean use.
- Higher recycling rates mean less landfill and incineration, reduced transport emissions, and more circular resource flow. This benefits air, soil, and marine environments. The Busan Facilities Corporation project proved that recycling is possible in previously unsorted areas, resulting in higher recycling rates and better-quality secondary materials.

## 2.3 Future Work

We are moving to innovate AI-based recycling systems, improve rates, and secure high-quality recycled raw materials. This requires finely tuned classification algorithms, robotic implementation, and logistics optimization through data.

Our next step is classifying plastic compositions. As global regulations mandate recycled plastic usage, we are developing high-speed plastic classification models using near-infrared sensors and vision systems.

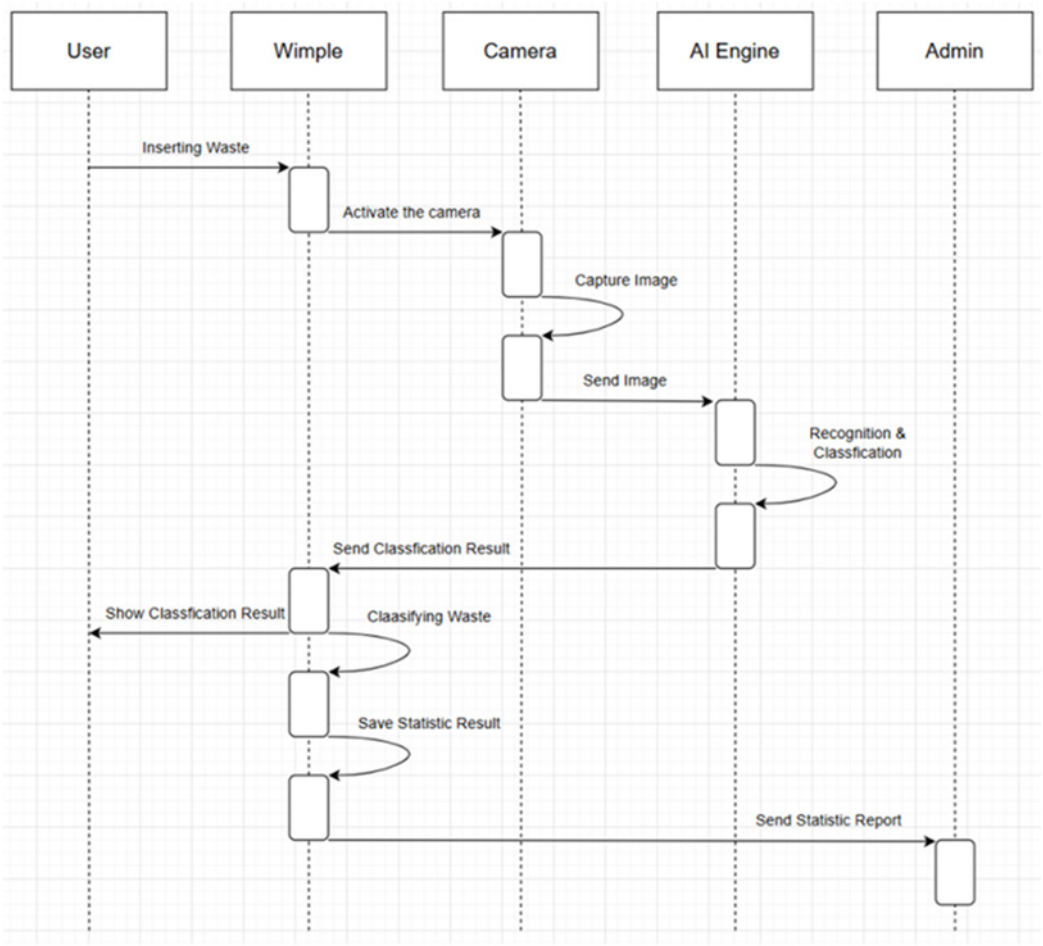
We continuously explore new sorting needs and develop models for finer classification. Our team is optimizing model accuracy and size. Entering new markets requires localized data. We partner with capable local firms for such market entries and continue advancing our generative AI and machine learning capabilities.

While AI is gaining traction in sectors like healthcare and automotive, its transformative potential in recycling remains underutilized. Canada, for instance, faced sorting and storage crises due to labour shortages—issues our AI helped resolve. The lack of AI application in recycling stems from a lack of domain-specific knowledge. We offer both technical expertise and field-tested know-how and are ready to collaborate with any partner globally. Beyond technological innovation, finding capable partners is another important next step. By working with more organizations that need us, we are one step closer to full circularity.

# 3 Use Case Requirements

- **REQ-01**: It is required to have datasets being designed flexibly for reuse across different contexts. The dataset from this solution is large-scale and structured for multiple uses as recycling containers are globally similar in material—glass, plastic, paper, metal—but vary in shape, size, and purpose.
- **REQ-02**: It is required to have effective integration and fast processing as classification is complex and requires multiple AI algorithms. The contamination detection must include removable, washable, or non-recyclable pollutants while sorting may differ by country (e.g., Canada separates aluminum vs. bi-metal cans).
- **REQ-03**: It is required to have a hardware system including recognition (for accurate detection), transport (to move and sort), and management (to leverage data and generate services).
- **REQ-04**: It is recommended to use WIMPLE-generated data as they can reduce costs by optimizing bin replacement timing, route planning for collection vehicles, and recycling logistics. Real-time carbon reduction metrics also support ESG reporting.

## 4 Sequence Diagram




## 5 References

[1] Jiacang Ho, JongHyuk Lee & HyoungSuk Kim. Wimplebin: an AI-based recycle bin for a better waste management Journal of Material Cycles and Waste Management 2025;27:584-596. (https://link.springer.com/10.1007/s10163-024-02145-9)

[2] Wimple M_Manual.pdf, "Wimple M_Manual.pdf," *Google Docs*, 2019. https://drive.google.com/file/d/1oPvgLn0NDl50VUJ2g_VHd-z1SZD9dH-l/view?usp=drive_link (accessed Jun. 09, 2025).

[3] Wimple Count_Manual.pdf, "Wimple Count_Manual.pdf," *Google Docs*, 2019. https://drive.google.com/file/d/1gBpDZj18KDaZv4UQZqvxMjV5ZVGdZAeE/view?usp=drive_link (accessed Jun. 09, 2025).

[4] "AI-Hub," *Aihub.or.kr*, 2023. https://www.aihub.or.kr/aihubdata/data/view.do?currMenu=115&topMenu=100&aihubDataSe=data&dataSetSn=71385 (accessed Jun. 09, 2025).

[5] *Bisco.or.kr*, 2025. https://www.bisco.or.kr/undershop/03_jag/jag01.asp (accessed Jun. 09, 2025).

[6] Busan – Cho Won-jin, "Dongseo University and Seore, Development and Verification of Advanced AI-Based Recyclable Waste Classifier," Seoul Economic Daily, Nov. 07, 2023. https://www.sedaily.com/NewsView/29X575U740 (accessed Jun. 09, 2025).

[7] *Bisco.or.kr*, 2025. https://www.bisco.or.kr/ (accessed Jun. 09, 2025).

# Use Case 7: LLM empowered smart hydrology

**Organization**: ZTE Corporation

**Country**: China

**Contact Person(S)**:

Primary: Liya Yuan, yuan.liya@zte.com.cn

Secondary: Wang Wang, wang.wang@zte.com.cn

## 1 Use Case Summary Table

| Item | Details |
|---|---|
| Category | Climate Change/Natural Disaster |
| Problem Addressed | Traditional hydrology struggles with slow data analysis, complex flood planning, and costly terrain surveys caused by manual-based methods. The integration of hydrology LLM facilitates the intelligent transformation and greatly improve the efficiency of water management system. |
| Key Aspects of Solution | Based on hydrology LLM, building a smart hydrological assistant delivering expertise and auto-reporting for water professionals; AI applications handling critical tasks like river modelling and engineering calculations; AI agents automating flood strategy design and emergency planning, optimizing entire operational chains. |
| Technology Keywords | Hydrology, LLM, multi-modal, agent, water governance |
| Data Availability | Private |
| Metadata (Type of Data) | text, visual |
| Model Training and Fine-Tuning | PEFT fine-tuning and instruction tuning based on the Nebula LLM with hydrology data |
| Testbeds or Pilot Deployments | N/A |

## 2 Use Case Description

### 2.1 Description

This case demonstrates a hydrology-specific LLM developed based on ZTE Nebula large model, designed to address core challenges in water conservancy. It focuses on optimizing

hydrological workflows, enhancing flood resilience, and enabling data-driven water resource governance.

Modern river basin management faces dual pressures from climate volatility and anthropogenic impacts. Traditional hydrological methods, often reliant on manual expertise, fragmented data, and labor-intensive field surveys, struggle to meet demands for precision and timeliness. Key challenges span real-time flood risk forecasting in complex river systems, high-resolution sediment transport modelling, adaptive water allocation balancing ecological and economic needs, and cost-effective infrastructure monitoring across diverse geographical conditions.

The hydrology large model enables an intelligent transformation and upgrade of water resource management based on its multi-modal capabilities, ultimately improving and liberating productivity in hydrology work. It serves as a hydrological knowledge assistant, providing domain-specific expertise and automated report generation services for water resource professionals. It also powers AI applications for critical scenarios including river evolution analysis, engineering computations, and hydrological testing with its visual analysis capabilities. Furthermore, the system implements intelligent hydrological agents based on the large model capable of flood prevention strategy formulation and emergency response planning, streamlining end-to-end workflows.

A standout application is the riverbed terrain prediction module, which plays a crucial role in flood control, channel maintenance, and water resource management. Traditional methods often rely on survey vessels to perform intensive and inefficient riverbed depth measurements. In this use case, the data from cross-sectional water depth measurements and high-frequency, high-resolution river monitoring enabled by satellite remote sensing technology can be analyzed by the large model and allows for highly efficient and accurate riverbed terrain predictions, demonstrating significant operational and economic advantages over conventional methods.

Through this model, we have achieved 90% accuracy in hydrology data queries, approximately 30% reduction in report writing time, over 20% improvement in thalweg line drawing efficiency and 80% accuracy in riverbed topography prediction.

Validated across diverse basin types, the system establishes a scalable paradigm for intelligent water governance. Its architecture bridges AI capabilities with hydrological expertise, setting a new standard for sustainable water management in the era of climate uncertainty.

Use Case Status: Operational

**Partners**: Changjiang Hydrological Bureau [3]

## 2.2 Benefits of use case

The introduction of the hydrology LLM drives digital and intelligent transformation in the water management sector. By leveraging AI innovations, it significantly enhances hydrological work efficiency and strengthens the intelligent management capabilities of water infrastructure. The model represents a technological breakthrough in modernizing critical infrastructure, promoting cutting-edge technological applications in water resource management.

By providing accurate hydrological data analysis and riverbed topography predictions, the model plays a crucial role in forecasting and mitigating water-related risks induced by climate change. It substantially improves flood prevention and water emergency management capacities, thereby

helping communities to both mitigate and adapt to the evolving environmental challenges. The advanced predictive capabilities enable proactive responses to increasingly complex climate-induced water system dynamics.

The model offers intelligent support for hydrological work, significantly enhancing disaster prevention and mitigation capabilities in urban and rural communities. Particularly in flood prevention, river management, and risk assessment, it effectively reduces water-related disaster risks and strengthens community resilience. By providing precise, real-time insights, the model empowers local authorities to make data-driven decisions that protect lives, infrastructure, and economic assets.

### 2.3 Future Work

In the forthcoming phase, we aim to create a more comprehensive knowledge base by integrating diverse data sources and advanced technologies. This will involve synthesizing meteorological data, Geographic Information Systems (GIS), hydrological historical datasets, and industry-specific historical records into a unified knowledge framework.

The strategic technological approach focuses on constructing an intelligent large model to optimize and support water management mechanism models. We will leverage cutting-edge technologies, including advanced sensing technologies, comprehensive network coverage, large language model capabilities, and cloud rendering technologies.

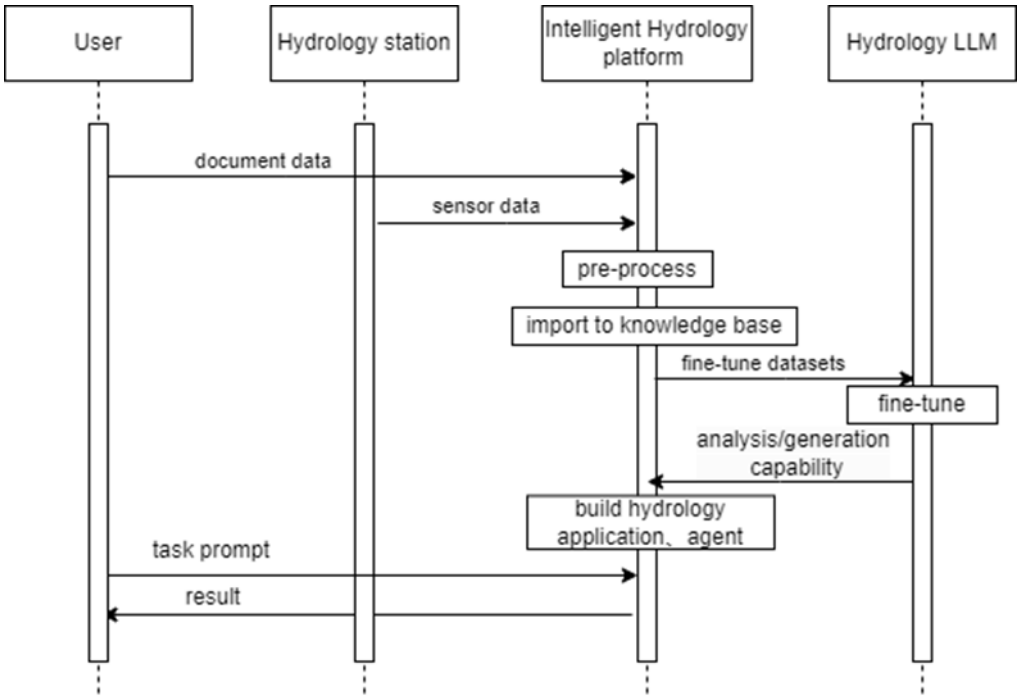
Our primary objectives encompass enhancing digital twin foundational capabilities and developing robust three-tier defense systems. This involves continuous exploration and implementation in critical scenarios such as video cascade control, data aggregation, rainfall radar infrastructure development, and communication guarantee in complex operational environments.

Ultimately, the goal is to provide comprehensive, multi-dimensional support for water resource management, driving technological innovation and operational efficiency in the water management sector. By integrating advanced computational techniques with deep domain expertise, we aim to transform how water resources are monitored, analyzed, and managed.

## 3 Use Case Requirements

- **REQ-01**: It is required that the hydrology station collects data like water level and flow data, rainfall and water quality data, soil and sediment data via different types of sensors and send to the intelligent hydrology platform as needed.
- **REQ-02**: It is required that the intelligent hydrology platform support data storage for further fine-tuning of the LLM.
- **REQ-03**: It is required to have multi-modal data processing for the LLM, including satellite remote sensing data, field survey inputs, and historical hydrological records.
- **REQ-04**: It is required to have cross-modal alignment for riverbed prediction, e.g., water depth measurements and satellite imagery.
- **REQ-05**: It is required to have 80%+ accuracy threshold for riverbed topography predictions.

## 4 Sequence Diagram




## 5 References

[1] The Yangtze River Hydrological Jiupai Large Language Model is Officially Released, Dec. 06, 2024. https://www.chinawater.com.cn/ly/zj/202412/t20241206_1059343.html (accessed Jun. 09, 2025).

[2] ZTE, "ZTE helps Changjiang Hydrological Bureau release the Jiupai Large Language Model and jointly promote the construction of hydrological digitalization, Dec. 13, 2024. https://mp.weixin.qq.com/s/36LUIkpEsXu7CnHBVFY9mg (accessed Jun. 09, 2025).

[3] Changjiang Hydrological Bureau. http://www.cjh.com.cn/ (accessed Jun. 09, 2025).

# Use Case 8: AI for green multi-cluster: Intelligent management towards green and low-carbon, large-scale multi-clusters

**Organization:** Institute of Artificial Intelligence (TeleAI), China Telecom

**Country**: China

**Contact Person(S)**:

Primary: Qizhen Weng, wengqzh@chinatelecom.cn

Secondary: Yuankai Fan, fanyk1@chinatelecom.cn

## 1 Use Case Summary Table

| Item | Details |
|---|---|
| Category | Climate Change/Natural Disaster |
| Problem Addressed | Using AI to efficiently managing computing resource and scheduling jobs across multiple large-scale AI clusters. This AI-driven approach aims to reduce energy consumption and carbon emissions, surpasses traditional rule-based multi-cluster management methods in adaptability and effectiveness, and is capable of explaining its decisions to humans via multi-modal generative models. |
| Key Aspects of Solution | • AI-powered job scheduling that improves efficiency through real-time multi-dimensional resource monitoring.<br>• Employing GPU multiplexing and shared resource scheduling algorithms to enhance cluster utilization.<br>• Cross-cluster coordination via multiple AI agents towards overall green and low-carbon goals.<br>• Employing multi-modal generative large models to provide interpretable explanations of scheduling decisions, enhancing transparency for end-users. |
| Technology Keywords | Job Scheduling, Multi-Cluster Management, Multi-Modal Generative Models, GPU Multiplexing, Kubernetes-based Cluster Coordination |
| Data Availability | Private; Public<br>1) https://github.com/alibaba/clusterdata [1]<br>2) https://github.com/InternLM/AcmeTrace [2]<br>3) https://github.com/ml-energy/zeus [3] |
| Metadata (Type of Data) | Text data (job resource requirements and performance logs, energy consumption metrics, energy source information, etc.) and images (cluster metrics represented through graphs) |

(continued)

| Item | Details |
|---|---|
| Model Training and Fine-Tuning | Models include reinforcement learning agents and large language models trained on extensive cluster resource and task operation data to learn optimal scheduling policies. Fine-tuning involves adapting models to dynamic multi-cluster environments, incorporating real-time monitoring data, and improving interpretability through multi-modal generative models. The Fragmentation Gradient Descent (FGD) algorithm is used for efficient GPU-sharing job scheduling, validated through simulation and real cluster deployments. |
| Testbeds or Pilot Deployments | • https://github.com/hkust-adsl/kubernetes-scheduler-simulator [4]<br>• https://github.com/qzweng/clusterdata/tree/master [5]<br>• https://www.usenix.org/conference/atc23/presentation/weng [6]<br>• https://www.usenix.org/conference/nsdi22/presentation/weng [7]<br>• https://sc20.supercomputing.org/proceedings/tech_paper/tech_paper_pages/pap211.html [8]<br>The links above contain information regarding deployments around scheduling GPU-sharing workloads, including novel measure of fragmentation to statistically quantify the extent of GPU fragmentation caused by different sources, Metis which is a general-purpose scheduler using deep reinforcement learning (RL) techniques, etc. |

# 2 Use Case Description

## 2.1 Description

The soaring popularity of generative AI models, such as ChatGPT [9], Midjourney [10], and DeepSeek [11], has enriched our lives and created a wealth of new employment opportunities. However, fueling AI requires substantial computational power and energy consumption, usually within AI clusters--data centers equipped with advanced AI processors, such as GPUs.

A recent report from the International Telecommunication Union (ITU) on international standards for AI and its environmental impact [12] highlights that energy consumption is driven primarily by the intensive processing required for training and inference of large AI models. These models are so energy-demanding that the computational power needed to support AI's expansion has been doubling around every 100 days since 2010 [13]. Beyond energy requirements, the vast quantities of fresh water needed to cool AI processors further strain local water resources and ecosystems [14]. This underscores the growing environmental footprint of AI and the urgent need to adopt efficient management practices for AI clusters.

**Intended Use:** Utilize artificial intelligence technology to achieve intelligent management and coordinated scheduling of multiple computing clusters, promote the green development of technology, enhance resource efficiency, and reduce energy consumption and carbon emissions.

**Problem to Solve:** With the increase in the number and scale of computing clusters, how to manage resources such as energy, computing power, and bandwidth in multiple clusters, and efficiently schedule and run a large number of computing tasks.

**Limitations of Existing Solutions:** Traditional multi-cluster management methods relying on manual experience and simple rules lack real-time performance and effectiveness when facing a multi-cluster environment with dynamically changing resource supply. They are unable to quickly identify the optimal resource allocation according to the actual situation.

Benefits and Drawbacks of the AI-based Approach:

1) It can adapt to environmental changes 24/7. Based on factors such as the number of running tasks, cluster computing power utilization, energy supply, and carbon emissions, it can flexibly adjust the resource scheduling scheme for multiple clusters to achieve the goal of green and low - carbon development.
2) It can perform performance modeling on computing tasks based on massive resource monitoring and task operation data, etc., thereby improving the overall computing performance.
3) It can use multi-modal generative large models to explain to humans the reasons behind scheduling decisions.

The proposed use case emphasizes the role of AI in enabling innovative and green and economical solutions for sustainable development. By proposing the requirements for intelligent low-carbon management, this system will boost the construction of data collection capabilities and low-carbon capabilities in industrial production. Additionally, by optimizing the utilization of infrastructure, the system enables limited infrastructure to support more tasks, indirectly reducing the need for new infrastructure construction and optimizing resource investment. This system complies with the goal of responsible consumption and production by optimizing the green utilization of resources throughout the AI lifecycle. It will promote green energy consumption patterns in computing clusters such as data centers and reduce waste and carbon emissions.

**Use Case Status:** The use case is part of a larger research project

Partners: N/A

## 2.2 Benefits of the use case

1) It can adapt to environmental changes 24/7. Based on factors such as the number of running tasks, cluster computing power utilization, energy supply, and carbon emissions, it can flexibly adjust the resource scheduling scheme for multiple clusters to achieve the goal of green and low - carbon development.
2) It can do performance modelling on computing tasks based on massive resource monitoring and task operation data, etc., thereby improving the overall computing performance.
3) It can use multi-modal generative large models to explain to humans the reasons behind scheduling decisions.

The proposed use case emphasizes the role of AI in enabling innovative and green and economical solutions for sustainable development. By proposing the requirements for intelligent low-carbon management, this system will boost the construction of data collection capabilities and low-carbon capabilities in industrial production. Additionally, by optimizing the utilization of infrastructure, the system enables limited infrastructure to support more tasks, indirectly reducing the need for new infrastructure construction and optimizing resource investment. This system complies with the goal of responsible consumption and production by optimizing the green utilization of resources throughout the AI lifecycle. It will promote green

energy consumption patterns in computing clusters such as data centers and reduce waste and carbon emissions.

## 2.3 Future Work

**Data Collection:** The dimension of data collection will be expanded to cover more energy and environmental parameters and detailed equipment operation information, so as to provide more comprehensive data support.

**Single-Node Resource Sharing:** Research will focus on developing more efficient resource-sharing algorithms to expand the scope of AI scheduling, including enhancements in GPU virtualization capabilities. The goal is to increase the utilization of GPU resources while minimizing idle waste, all while ensuring that task performance is maintained. Drawing from our operational experience at the Institute of AI, research teams across domains such as Computer Vision (CV), Natural Language Processing (NLP), and Speech have already virtualized a portion of their GPU resources. This GPU-sharing approach has led to an increase of at least 150% in the availability of GPUs for prototype development and debugging, without the need for additional processor purchases.

**Single-Cluster Resource Scheduling:** Monitoring and task data will be regarded as the state space, and scheduling optimization strategies will be regarded as the action space. Technologies such as reinforcement learning and large language models will be used to establish intelligent scheduling strategies, making them more suitable for complex and variable task requirements and further improving the resource utilization efficiency within a single cluster. The de facto standard for cluster coordination is Kubernetes, which is widely adopted by leading AI companies such as OpenAI [12]. Building on Kubernetes, we have open-sourced a scheduling algorithm known as the Fragmentation Gradient Descent (FGD) policy [13]. This algorithm enables highly efficient scheduling decisions within hundreds of milliseconds on clusters comprising over 1,200 servers. Simulations have demonstrated that FGD outperforms traditional bin-packing algorithms in GPU-sharing scenarios, reducing unallocated GPUs by up to 49% [14].

**Multi-Cluster Collaborative Scheduling:** With the help of technologies such as multi-agent large models, a more intelligent cross-cluster collaboration mechanism will be constructed. According to the real-time status of each cluster and task priorities, the optimal cross-cluster resource allocation will be achieved, promoting the achievement of overall green and low-carbon goals.

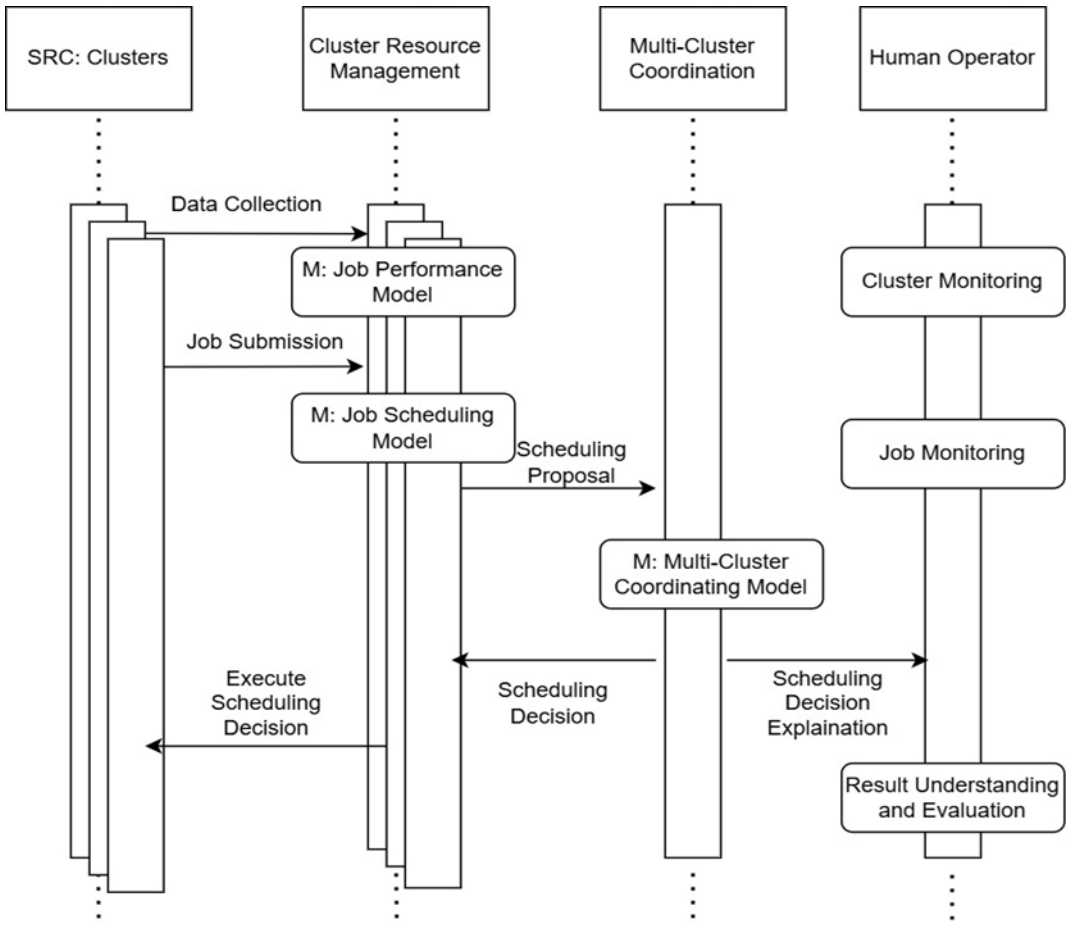
**Cluster Scheduling Result Interpretability:** Multi-modal generative large models will be utilized to present the scheduling decision-making process to users in an understandable form, enhancing trust in artificial intelligence scheduling results and facilitating subsequent strategy adjustment and optimization.

# 3 Use Case Requirements

- **REQ-01**: It is required to monitor multi-dimensional resource usage, profile typical job resource requirements, and collect environmental and infrastructural data from AI clusters.
- **REQ-02**: It is required to develop data-driven scheduling algorithms capable of adapting to dynamic resource availability, and conduct a preliminary evaluation of their performance using a cluster simulator.

- **REQ-03**: It is recommended to enable resource sharing to maximize hardware utilization.
- **REQ-04**: It is recommended to coordinate multiple clusters with intelligent agents to make wise scheduling decisions.
- **REQ-05**: It is recommended to use multi-modal generative models to explain scheduling decisions to human users.

## 4 Sequence Diagram




## 5 References

[1] "Alibaba Cluster Trace Program," *GitHub*, Feb. 04, 2023. https://github.com/alibaba/clusterdata

[2] InternLM, "GitHub – InternLM/AcmeTrace," *GitHub*, 2023. https://github.com/InternLM/AcmeTrace (accessed Jun. 09, 2025).

[3] ml-energy, "GitHub – ml-energy/zeus: Measure and optimize the energy consumption of your AI applications!," *GitHub*, May 17, 2025. https://github.com/ml-energy/zeus (accessed Jun. 09, 2025).

[4] hkust-adsl, "GitHub – hkust-adsl/kubernetes-scheduler-simulator: Kubernetes Scheduler Simulator," *GitHub*, May 11, 2023. https://github.com/hkust-adsl/kubernetes-scheduler-simulator.

[5] qzweng, "GitHub – qzweng/clusterdata: cluster data collected from production clusters in Alibaba for cluster management research," *GitHub*, 2021. https://github.com/qzweng/clusterdata/tree/master (accessed Jun. 09, 2025).

[6] W. Gropp, Torsten Hoefler, Rajeev Thakur, and Ewing Lusk, *Using advanced MPI modern features of the Message-Passing Interface*. Cambridge, Mass. The Mit Press, 2014.

[7] Q. Weng *et al.*, "{MLaaS} in the Wild: Workload Analysis and Scheduling in {Large-Scale} Heterogeneous {GPU} Clusters," *www.usenix.org*, 2022. https://www.usenix.org/conference/nsdi22/presentation/weng

[8] "Metis: Learning to Schedule Long-Running Applications in Shared Container Clusters at Scale," *Supercomputing.org*, 2025. https://sc20.supercomputing.org/proceedings/tech_paper/tech_paper_pages/pap211.html (accessed Jun. 09, 2025).

[9] OpenAI, "ChatGPT," *ChatGPT*, 2025. https://chatgpt.com/

[10] Midjourney, "Midjourney," *Midjourney*, 2024. https://www.midjourney.com/

[11] deepseek-ai, "GitHub – deepseek-ai/DeepSeek-V3," *GitHub*, 2024. https://github.com/deepseek-ai/DeepSeek-V3

[12] Wang L., Weng Q., Wang W., Chen C., Li B. (2020). Metis: Learning to Schedule Long-Running Applications in Shared Container Clusters at Scale. *SC*. Retrieved from https://sc20.supercomputing.org/proceedings/tech_paper/tech_paper_pages/pap211.html

[13] Weng Q., Xiao W., Yu Y., etc. (2022). MLaaS in the Wild: Workload Analysis and Scheduling in Large-Scale Heterogeneous GPU Clusters. *USENIX NSDI*. Retrieved from https://www.usenix.org/conference/nsdi22/presentation/weng

[14] Weng Q., Yang L., Yu Y., etc. (2023). Beware of Fragmentation: Scheduling GPU-Sharing Workloads with Fragmentation Gradient Descent. *USENIX ATC*. Retrieved from https://www.usenix.org/conference/atc23/presentation/weng

# Use Case 9: Combating wildlife trade with an advanced AI solution: a collaboration between Baidu PaddlePaddle and IFAW

**Organization**: Baidu

**Country**: China

**Contact Person(s)**:

Primary: Tianmin Long, intlcomm@baidu.com

Secondary: Yang Liu, liuyang145@baidu.com

Additional Contacts: Cathy Fu, fuduanling@baidu.com

## 1 Use Case Summary Table

| Item | Details |
|---|---|
| Category | Climate Change/Natural Disaster |
| Problem Addressed | AI image solution to fight against wildlife cybercrime |
| Key Aspects of Solution | Large Vision Model, Semi-supervised Learning, Zero-code Platform |
| Technology Keywords | Image Identification, Artificial Intelligence |
| Data Availability | Public |
| Metadata (Type of Data) | Text, Visual (Images) |
| Model Training and Fine-Tuning | Semi-supervised Learning using PaddlePaddle with IFAW image data |
| Testbeds or Pilot Deployments | Baidu AI Studio, Demo Video |
| Code Repositories | N/A |

## 2 Use Case Description

### 2.1 Description

According to the IUCN Red List, nearly 4,000 species are critically endangered, and the illegal wildlife trade is a major contributor to this biodiversity crisis. With much of the trade moving

online, conservation groups have been seeking innovative technological solutions to combat wildlife cybercrime.

The International Fund for Animal Welfare (IFAW), a leading nonprofit in animal conservation, initially relied on manual monitoring to identify online illegal wildlife trade. However, this labor-intensive process was inefficient and prone to errors. In 2019, IFAW partnered with Baidu to launch AI Guardian 1.0, powered by Baidu's deep learning platform PaddlePaddle. The system used image recognition to detect trade in elephant, pangolin, and tiger products.

By mid-2024, the tool had screened over 360,000 images and led to the removal of 7,853 illegal postings. In November 2024, AI Guardian 2.0 was launched to target the growing market of live exotic pet sales. Trained on a decade's worth of IFAW image data, the updated model can detect 34 species and boasts an 86% accuracy rate. A zero-code interface enables non-engineers to customize and deploy the model for real-time monitoring of online platforms.

Use Case Status: Operational

**Category:** Endangered Species Protection

## 2.2 Benefits of use case

The platform detects illegal wildlife products and support legal enforcement and thereby promoting sustainable practices, reducing biodiversity loss, and protecting marine and terrestrial ecosystems.

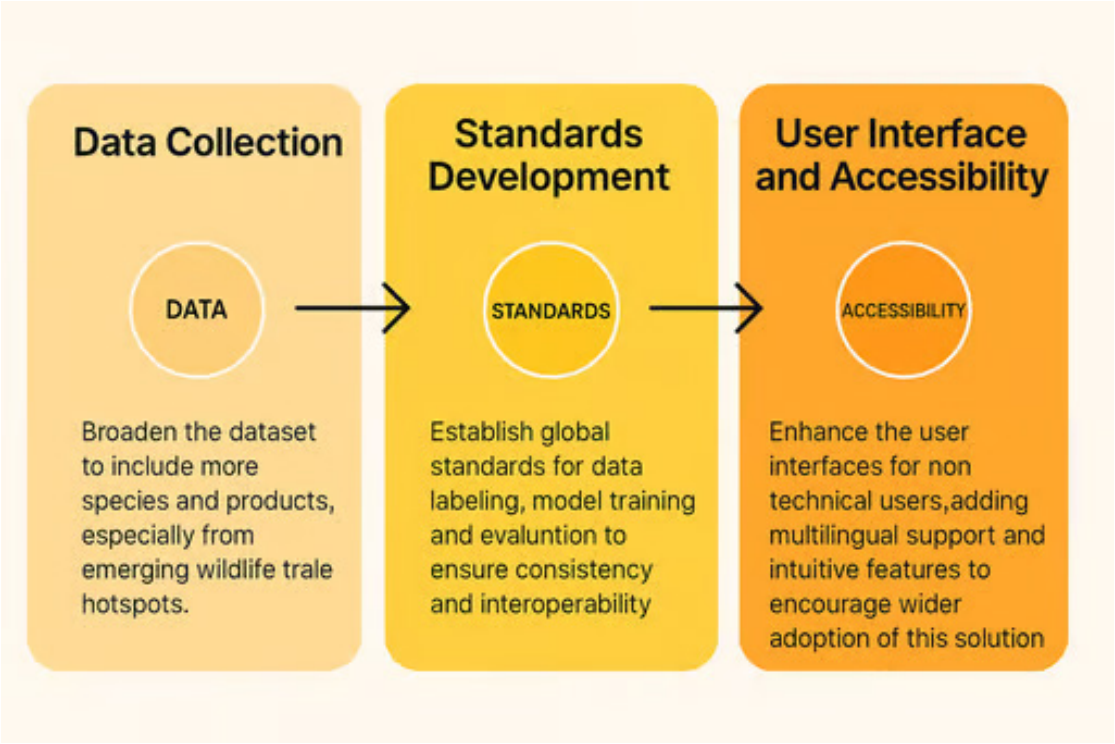
## 2.3 Future Work

- Data Expansion: Include more endangered species and items from different regions.
- Capacity Building: Launch training modules for enforcement agencies and NGOs.
- Regional Pilots: Expand AI Guardian deployment in Southeast Asia and beyond.
- Interface Improvements: Add multilingual support and simplify model deployment tools.
- Open Source Collaboration: Encourage developers to create custom biodiversity protection tools using PaddlePaddle.

# 3 Use Case Requirements

- **REQ-01**: It is critical to address the scarcity of annotated image data for illegal wildlife products.
- **REQ-02**: It is critical to handle class imbalance, as negative samples far outnumber positive ones.
- **REQ-03**: It is critical to accurately differentiate between real wildlife products and synthetic imitations.
- **REQ-04**: It is critical to detect altered features in processed wildlife goods (e.g., carved or dyed).
- **REQ-05**: It is critical to enable fast scanning and flagging across digital platforms to prevent illegal sales.

## 4 Sequence Diagram ( Tobe updated)




## 5 References

[1] https://aistudio.baidu.com/projectdetail/paddlex/6809846

[2] https://youtu.be/4XxgdIrDS_U?si=UO9IgbmbARmoYdqb

[3] https://mp.weixin.qq.com/s/KFJ2Mf6HA2ajyTQu7t-HpA

[4] T. Smith et al., "AI-Powered Wildlife Crime Detection," *Journal of Conservation AI*, vol. 11, no. 2, 2023.

[5] A. Liu et al., "Deep Learning for Wildlife Protection," *Proc. of the Intl. Conf. on AI & Biodiversity*, 2022.

# Use Case 10: Quartz Solar AI Cloudcasting - predicting cloud movements through satellite imagery

Open Climate Fix

**Organization**: Open Climate Fix

**Country**: United Kingdom

**Contact Person(s)**:

**Primary Contact**: Dan Travers – dan@openclimatefix.org

**Secondary Contact**: Kasia Krasucka – kasia@openclimatefix.org

**Additional Contact**: James Fulton, PhD – james@openclimatefix.org

## 1 Use Case Summary Table

| Item | Details |
|---|---|
| Category | Climate Change/Natural Disaster |
| Problem Addressed | Inaccurate solar forecasts due to unpredictable cloud cover, creating challenges in renewable energy integration and grid reliability. |
| Key Aspects of Solution | – AI model predicts future satellite imagery up to 3 hours ahead<br>– Secondary AI model forecasts solar PV generation based on cloud predictions |
| Technology Keywords | machine learning, neural network, computer vision |
| Data Availability | EUMETSAT public satellite data:[4] |
| Metadata (Type of Data) | Visual, video, satellite |
| Model Training and Fine-Tuning | – simVP architecture to forecast satellite images<br>– Two models: one for cloud prediction, one for PV power forecasting |
| Testbeds or Pilot Deployments | Open Climate Fix. n.d. *Cloudcasting App Repository*.[4] |
| Code Repositories | Open Climate Fix. n.d. *SAT Pred Repository*. [6]<br><br>Alan Turing Institute. n.d. *Cloudcasting Project Repository*.[7] |

# 2 Use Case Description

## 2.1 Description

The transition to renewable energy has made weather a critical factor in grid management. Cloud cover remains a primary cause of solar energy forecast inaccuracy. Open Climate Fix developed an open-source AI system that forecasts cloud movements using satellite imagery, improving short-term solar power prediction.

This innovation has increased solar forecast accuracy by 5% at a 2-hour horizon—double the accuracy gain seen with traditional weather services. The AI model is now being integrated into the Quartz Solar platform for the UK, in partnership with the National Energy System Operator (NESO), and is expected to reduce carbon emissions by over 100,000 tonnes annually.

The system uses the simVP architecture to predict future satellite images across 11 spectral channels and another model to convert those predictions into solar PV power forecasts. This method outperforms current ECMWF and Met Office models and is designed to scale globally to aid electricity grid operators, solar providers, and policymakers.

**Use Case Status:** Pilot integration planned for Summer 2025 in the UK

Partners

**Alan Turing Institute** – Research collaboration during the early development phase.

## 2.2 Benefits of use case

- Improved forecasts support solar energy integration and smart grid management.
- Demonstrates cutting-edge AI and infrastructure support for smart energy systems.
- Reduces reliance on fossil fuels and supports grid decarbonization efforts.

## 2.3 Future Work

The development roadmap includes:

- Stage 1: Launch in the UK with Quartz Solar (Summer 2025)
- Stage 2: Expand to India and the Netherlands (2025–2026)
- Stage 3: Build and release a global, open-source cloudcasting service

Target adoption strategy includes:

- Large grids with forecasting needs (e.g., US, Japan, India)
- High-solar penetration markets (e.g., Brazil, Southeast Asia)
- Emerging and decentralized grids (e.g., Sub-Saharan Africa)

Deployment will combine direct services and open-source tools to foster global innovation.

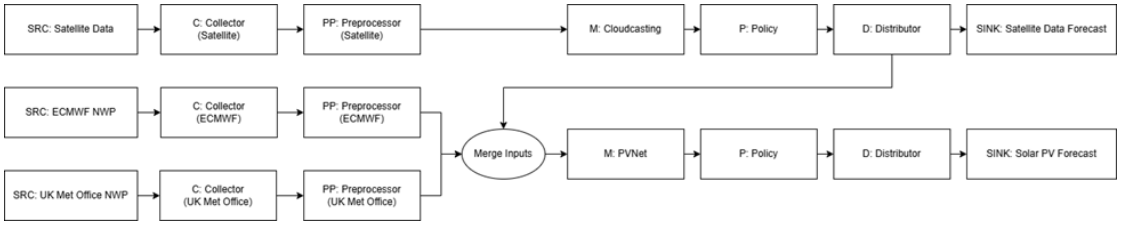
# 3 Use Case Requirements

- **REQ-01**: The AI model must accurately predict future satellite images across 11 spectral channels.

- **REQ-02**: The AI model must measurably improve solar PV forecasts over using satellite data alone.
- **REQ-03**: The system must run efficiently and use minimal compute resources for easy deployment and real-time operation.

## 4 Sequence Diagram

Sequence Diagram link




## 5 References

[1] Travers, Dan, et al. 2024. "Cloudcasting for Solar Forecasting." *Climate Change AI – ICLR 2024*. https://github.com/openclimatefix/cloudcasting-app.

[2] Travers, Dan. 2023. "ICEM 2023 Presentation." https://github.com/openclimatefix/sat_pred.

[3] Travers, Dan. 2024. "ICLR 2024 Conference Talk." https://github.com/alan-turing-institute/cloudcasting.

[4] Google Cloud Platform. n.d. *EUMETSAT SEVIRI RSS HRV UK Dataset on BigQuery Public Data*. https://console.cloud.google.com/marketplace/product/bigquery-public-data/eumetsat-seviri-rss-hrv-uk.

[5] Open Climate Fix. n.d. *Cloudcasting App Repository*. https://github.com/openclimatefix/cloudcasting-app.

[6] Open Climate Fix. n.d. *SAT Pred Repository*. https://github.com/openclimatefix/sat_pred.

[7] Alan Turing Institute. n.d. *Cloudcasting Project Repository*. https://github.com/alan-turing-institute/cloudcasting.

# Use Case 11: AI biodiversity monitoring

**Organization**: SBER

**Country**: Russia

**Contact Person(s)**:

**Primary Contact**: Iuliia Zemtsova – YMZemtsova@sberbank.ru

**Secondary Contact**: Oleg Artyugin – oyartyugin@sberbank.ru

Other Contacts:

Konstantin Gongalskii – Gongalsky@sev-in.ru

Jose A Hernandez-Blanco – j.a.hernandez.blanco@yandex.ru

Denis Malikov – zoolog.22@yandex.ru

## 1 Use Case Summary Table

| Item | Details |
|---|---|
| Category | Climate Change/Natural Disaster |
| Problem Addressed | AI biodiversity monitoring |
| Key Aspects of Solution | Use of AI for optimized camera trap monitoring to classify 25 native species in the Russian Far East and Altai |
| Technology Keywords | machine learning, computer vision, conservation, ecology, camera traps, wildlife, MegaDetector |
| Data Availability | Private (Plans to open-source repository in 2025) |
| Metadata | Photo |
| Model Training | 73,687 images (60/20/20 train/val/test split) using EVA2 model (F1-macro: 0.93) |
| Pilot Deployment | https://biodiversity.ai4good.ru/ |
| Code Repositories | Not publicly available |

## 2 Use Case Description

### 2.1 Description

Camera traps, deployed globally for ecological monitoring, capture vast amounts of imagery triggered by motion or on a set schedule. These images are traditionally reviewed manually, which is labor-intensive and time-consuming—often cluttered with irrelevant (empty or human/vehicle) images.

This project integrates AI to automate species identification from camera trap images. Our model suite includes a detector (MegaDetector) followed by a classifier capable of distinguishing 25 native species, thereby reducing manual effort and accelerating biodiversity assessments.

### Partners

- Severtsov Institute of Ecology and Evolution, RAS[4]
- Saylyugemsky National Park[5]

**Partner Roles:**

- Data provision
- Piloting implementation

Providing domain expertise

## 2.2 Benefits of use case

To conserve ecosystems, AI is used to monitor species and their food chains by automating image classification, organizing animal data, and enhancing expert workflows.

## 2.3 Future Work

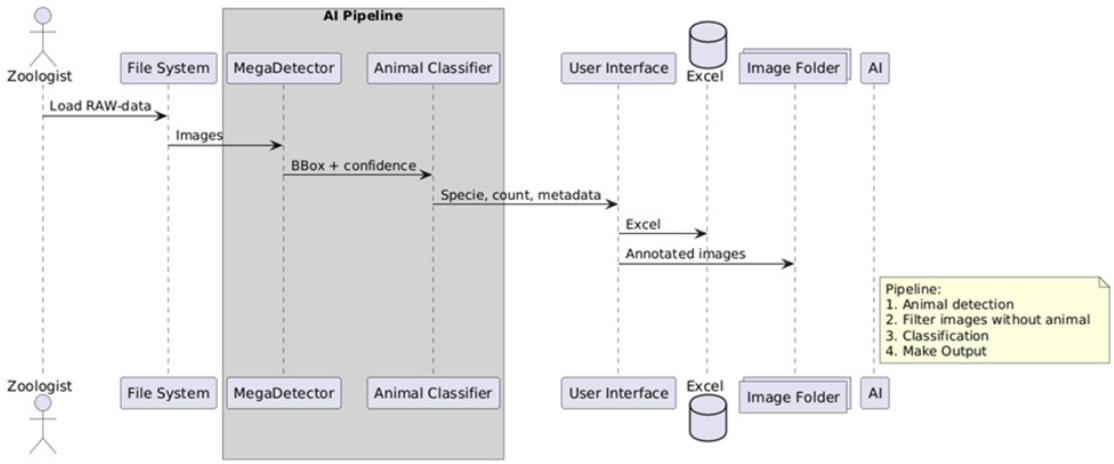
- Expand to additional species beyond the initial 25
- Incorporate new data modalities (e.g., audio, video)

Develop an interactive AI agent to assist zoologists

# 3 Use Case Requirements

- **REQ-01:** Classification accuracy over 90%
- **REQ-02**: Ability to classify multiple species in one image
- **REQ-03**: Ability to distinguish similar species
- **REQ-04**: Output must follow standard formats (e.g., EcoAssist)
- **REQ-05**: Must support local execution (offline capability)

# 4 Sequence Diagram



## 5 References

[1] LILA.science. n.d. *Dataset Benchmark*. https://lila.science.

[2] MegaDetector. n.d. *MegaDetector GitHub Repository*. https://github.com/agentmorris/MegaDetector.

[3] Biodiversity AI Pilot. n.d. https://biodiversity.ai4good.ru/.

[4] Severtsov Institute of Ecology and Evolution, RAS. n.d. https://sev-in.ru/.

[5] Saylyugemsky National Park. n.d. https://sailugem.ru/.

# Use Case 12: AI for Weather Prediction for Africa (Project Umoya)

**Organization:** Ishango.ai

**Country**: United Kingdom

**Contact Person**:

**Primary**: Eunice Baguma Ball, eunice@ishango.ai | +44 7857 449119

## 1 Use Case Summary Table

| Item | Details |
|---|---|
| Category | Climate Change/Natural Disaster |
| Problem Addressed | Lack of localized AI-based weather prediction (AIWP) models and expertise for the African continent |
| Key Aspects of Solution | – Community building among African AI experts<br>– Local dataset collection and curation<br>– Fine-tuning AIWP models like GraphCast for Africa<br>– Supporting local weather services and institutions |
| Technology Keywords | AIWP, GraphCast, machine learning, climate adaptation, reanalysis data, low-cost sensors, weather forecasting |
| Data Availability | No public or private access yet (future datasets planned) |
| Metadata (Type of Data) | Reanalyses, Automated Weather Station data, Trained model parameters |
| Model Training and Fine-Tuning | Models such as GraphCast and Aurora will be trained or fine-tuned using African weather datasets including ERA5 and station data |
| Testbeds or Pilot Deployments | Proof-of-concept in progress – community training, data sourcing, and local forecasting capability building |

## 2 Use Case Description

### 2.1 Description

Project **Umoya** (isiZulu for "wind") addresses a critical gap in African weather forecasting by building local capacity in AI for Weather Prediction (AIWP). While models like DeepMind's **GraphCast** show excellent performance globally, Africa lacks locally tailored solutions due to limited technical expertise, infrastructure, and data availability.

Objectives:

- Create Africa's first open-source AIWP model, built by Africans for Africa

- Train a cohort of African data scientists in AIWP modeling and forecasting
- Source and curate data from public datasets (ERA5, WMO AWS) and unique local sources (WASA, AfriNet, WNS)
- Support local organizations and government bodies to integrate AIWP tools into weather-sensitive sectors

**Approach**:

- Use existing AIWP model weights from GraphCast/Aurora
- Incorporate station-level and satellite data
- Leverage low-cost infrastructure and local community knowledge
- Promote open access and reproducibility of tools and insights

Use Case Status: Research and Product Development Phase

Partners: African Institute for Mathematical Sciences (AIMS)

## 2.2 Benefits of use case

1. weather prediction and local climate model for Africa. Brings benefits to weather sensitive sectors such as agriculture
2. Incorporating local community knowledge and supporting local organizations and governments.
3. promoting open access, tools and insights.

## 2.3 Future Work

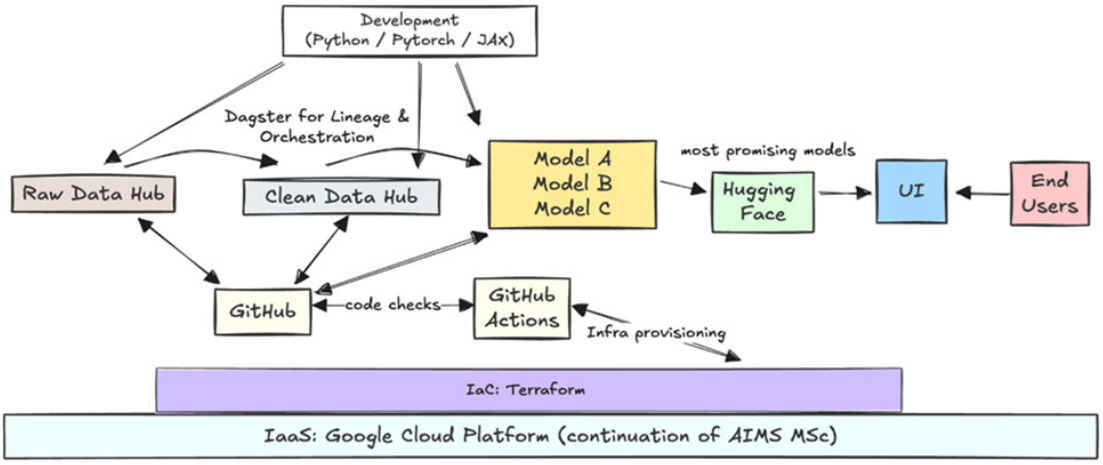
- Expand curated datasets across additional African regions and climatic zones
- Develop and validate proof-of-concept AIWP models for deployment in local sectors (agriculture, disaster risk, energy)
- Publish open-source models and data pipelines

Host regional training workshops and establish mentorship programs for African AIWP researchers

# 3 Use Case Requirements

- **REQ-01**: It is critical that models be fine-tuned or developed specifically for African weather conditions
- **REQ-02**: It is expected that curated datasets include both reanalysis data and ground-based AWS inputs
- **REQ-03**: It is of added value to incorporate low-cost sensors and community-sourced data
- **REQ-04**: It is expected that AIWP models can run efficiently on limited computational infrastructure
- **REQ-05**: It is critical to establish a sustainable pipeline for local expert training and community participation

## 4 Sequence Diagram




## 5 References

[1] DeepMind (2022). *GraphCast: Learning skillful medium-range global weather forecasting*. Retrieved from https://www.deepmind.com/blog/using-ai-to-accurately-forecast-weather

[2] ERA5 – European Centre for Medium-Range Weather Forecasts (ECMWF)

[3] WASA Project – https://www.wasaproject.info

[4] AfriNet – https://www.africlimate.ai/projects/afrinet

[5] WNS – Agricultural Research Council (ARC), Weather Station Network

# Use Case 13: Analysing Impact of Regional Weather on PLC-based Data Transmission in Malaysia, and Nigeria.

**Organization:** Wireless & Photonic Networks (WiPNet) Research Center, Universiti Putra Malaysia (UPM)

**Country:** Malaysia and Nigeria

**Contact Person(s)**:

Muhammad Adamu, m.adamu.ce@gmail.com

Prof. Dr. Ir. Aduwati Sali, aduwati@upm.edu.my

## 1 Use Case Summary Table

| Item | Details |
|---|---|
| Category | Climate Change/Natural Disaster |
| Problem Addressed | The changing of climate threatens data communication by disrupting weather patterns and shapes which make it becoming more adverse and extreme thereby causing potential impact on the performance of data transmission rate (Throughput). |
| Key Aspects of Solution | Investigating significant weather events that potentially influence Power Line Communication(PLC)-based transmission systems regarding Throughput (data rate). |
| Technology Keywords | Utilising traditional Machine Learning(ML)approach of 'Neural Network Regularization(NNR)to analyse the potential impact of events on PLC-based data transmission with respect to Throughput performance. |
| Data Availability | Investigating industrial practices to assess compliance with relevant standards (International Telecommunication Union(ITU)-T G.996X, IEEE 1901). |
| Metadata | Related projects will be used for comparison. |
| Model Training | Power Line Communication (PLC), Throughput, Weather Events, Machine Learning (ML), Neural Network Regression (NNR). |
| Pilot Deployment | Public |
| Code Repositories | Text (datasets, and scripts (Python, MATLAB)) |

# 2 Use Case Description

## 2.1 Description

The change in climate characteristics disrupts weather shapes and patterns thereby making it becoming more extreme and adverse. Climate-inducing terrible weather events have potential impact on several activities such as in agriculture, transportation, healthcare, transmission, and so on. Extreme weather events (like temperature, rain, thunderstorms, humidity, etc.) potentially affect the quality of data transmission over communication networks. The PLC is a wireline-based communication technology that support simultaneous transmission of electricity and communication data power lines using the traditional power cabling infrastructures. Climate-inducing extreme weather events cause several data transmission challenges including; attenuation, interference, noise, surge, lost, disruption, distortion, and so on. This project investigates regional extreme weather events and analyses their potential impact on PLC-supported power lines with respect to data transmission in Peninsular Malaysia and Nigeria. The project developed an NNR-based ML model accepts weather and predict its impact.

The objectives of the project are;

- Investigate extreme weather events that potentially impact PLC-based power lines in Malaysia and Nigeria.
- Analyse potential impact of the investigated extreme weather events on PLC-based power lines with respect to data transmission rate.
- Assess the performance of data rate over regional PLC-based power lines in Malaysia and Nigeria with reference to ITU-T G.9960 (2008-2023) recommendations on data performance of wireline-based technologies like telephone, Ethernet, and PLC system. Some related projects are also used in the assessment for comparison.

This project provides regional weather-based guides to relevant stakeholders (industries, regulatory bodies, users/consumers, government, etc.) for effective and responsible implementation, utilization, maintenance, and optimisation of PLC systems.

**Use Case Status:** This use case is in the research and prototyping stage, with real-world data collection and model development already conducted for Malaysia and Nigeria.

**Partners:** Power transmission industries are the major partners in this project. In Nigeria, the Transmission Company of Nigeria (TCN) is utilising PLC network system over their power lines. The system currently undergoing overhaul to replace traditional power cables with high-data cables due to technical and environmental (weather) challenges. Likewise in Malaysia, the Tenaga Nasional Berhad (TNB) (a major utility company for electricity) is also assessing similar challenges, and therefore is replacing copper lines with fiber optic lines for high data delivery and resilient to tropical weather events of Malaysia.

## 2.2 Benefits of use case

This project is significant in the utilisation of PLC-based resources that are essential for data transmission. The analysis in this project ensures responsible usage of data operating under climate-inducing harsh weather. It provides weather-based references to power stakeholders on deployment, maintenance, and optimisation of extremely-weather-exposed PLC systems in Malaysia and Nigeria. The project is transferable to analyse other regions (beyond Malaysia and Nigeria) and communication technologies (like cellular networks, wireless, satellite, radio/

TV broadcasting system, IoT, etc.); upon the provision of additional weather and particular transmission requirements.

## 2.3 Future Work

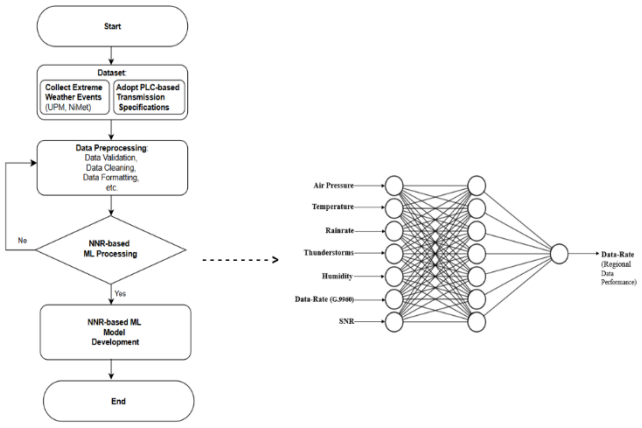
- Data Collection: The project utilises live weather events, therefore continuous collaboration with relevant meteorological agencies is necessary for continued capturing of high-quality weather events. The collaboration ensures that the model is continuously train using up-to-date and consistence data. Partnership with regulatory bodies can be established to strengthen the model's generality.
- Model's Development: In the project, an NNR-based ML model analyses the impact of weather events on PLC-based system regarding data rate in Malaysia and Nigeria. The model's performance can further be evaluated with other traditional ML models (such as Polynomial Regression, Support Vector Regression, Random Forest Regression, Gradient Boosting Regression, k-NN Regression, Gaussian Process Regression, etc.). The ML training and hyperparameters can be maximised and optimised respectively to improve accuracy and efficiency.
- Use Case Difference: The NNR-based ML model is developed to analyse impact of extreme weather events on PLC system. The PLC is a wireline-based technology that support power and data transmission using similar power cabling infrastructure. By providing additional weather and specific transmission requirements, the model's capabilities can be extended to analyse other regions and communication technologies.

# 3 Use Case Requirements

- **REQ-01**: The Use Case domain is to analyse impact of climate change-inducing extreme weather on data transmission of PLC-supported power lines in Malaysia and Nigeria. Other communication technologies and regions are not captured.
- **REQ-02**: It is critical for power industries to deploy or maintain PLC systems based on the ITU-T G.9960 recommended architecture and performance. Most regional PLC-based power systems use different parameters. It is also important maintain partnership with stakeholders (power industries, weather monitoring agencies, and regulatory bodies) for effective prediction of current and future data performances.
- **REQ-03**: It is critical to continue sourcing for up-to-date weather data, as climate change is continuously disrupting weather patterns. Datasets and scripts utilised in the project are publicly available at https://github.com/adams714040/ITU-Project.git.[1]
- **REQ-04**: It is critical to explore other ML-based regression models that are suitable in analysing weather patterns and impact such as polynomial regression, support vector regression, random forest, Gradient Boosting, k-NN, gaussian process, etc.) for accuracy, efficiency, and reliability.
- **REQ-05**: It is critical to align prediction carried out in this project with other relevant standards like IEEE 1901 (2010).

# 4 Sequence Diagram

The figure below shows development process of the NNR-based ML model. Dataset is collected and validated. It is then preprocess, and process onto ML for training. The ML processing is computed using 2-layer structure; where the first layer takes the input (dataset), and the second layerenables the NNR to learns and trains the data for prediction (output).




## 5 References


[1] https://github.com/adams714040/ITU-Project.git

[2] P. Chebotayev1, D. Urazayev, I Pospelova, A. Karassenko, D. Zykov, A. Shelupanov. "Analysis of the level of SNR in the PLC channel based on the results of mathematical and physical modelling". Journal of Physics: 2019 DOI:10.1088/1742-6596/1145/1/012013

[3] Konark Sharman, Lalit Mohan Saini. "Power-Line Communications for Smart Grid: Progress, Challenges, Opportunities and Status". Renewable and Sustainable Energy Reviews. Vol. 67, January 2017, Pages 704-751. https://doi.org/10.1016/j.rser.2016.09.019

[4] Peter L. Watson, Aaron Spaulding, Marika Koukoula, Emmanouil Anagnostou. "Improved quantitative prediction of power outages caused by extreme weather events". Weather and Climate Extremes. Volume 37,2022, https://doi.org/10.1016/j.wace.2022.100487

[5] Yue Wang, Lu Chen, Meng Yao, Xinyu Li. "Evaluating weather influences on transmission line failure rate based on scarce fault records via a bi-layer clustering technique". IET Generation, Transmission Distribution. ISSN 1751-8687 E-First on 7th November 2019 doi: 10.1049/iet-gtd.2019.0551

[6] Tolulope David Makanju, Oluwole John Famoriji, and Thokozani Shongwe. "Machine Learning Approaches for Power System Parameters Prediction: A Systematic Review". IEEE Access 2017. Department of Electrical and Electronic Engineering Technology, University of Johannesburg, P.O. Box 524, Auckland Park, 2006, Johannesburg, South Africa. DOI 10.1109/ACCESS.2024.3397676

[7] Muhammad Rahman, Mohammad Azazur Rahman, Muhammad Mostafa Amir Faisal, Syed Zahidur Rashid. "Feasibility of Power Line Communication in Bangladesh and Analyzing the Power Spectral Density of Colored Background Noise". Conference: International Conference on Industrial Engineering and Operations Management. Washington DC, USA, Septem- ber 27-29, 2018

[8] Wenzhe Bi, Li Tian, Chao Li, Zhen Ma, Haiyang Pan. Wind-induced failure analysis of a transmission tower-line system with long-term measured data and orientation effect. Reliability Engineering and System Safety 229 (2023) 108875. https://doi.org/10.1016/j.ress.2022.108875

[9] Abdul Mannan, D.K. Saxena, Mahroosh Banday. "A Study on Power Line Communication". International Journal of Scientific and Research Publications, Volume 4, Issue 7, July 2014 ISSN 2250-3153

[10] Olga Kondrateva, Ekaterina Myasnikova, OlegOleg Loktionov. "Analysis of the Climatic Factors Influence on the Overhead Transmission Lines Reliability". Environmental and Climate Technologies 2020, vol. 24, no. 3, pp. 201–214 https://doi.org/10.2478/rtuect-2020-0097

# Use Case 14: Remote Visual Tracking Using Semantic Scene Understanding and GenAI-based Predictive Rendering Under Disaster Situation

**Organization:** Tata Consultancy Services

**Country**: India

**Contact Person(S)**:

**Primary**: Abhijan Bhattacharyya, Abhijan.bhattacharyya@tcs.com

**Secondary**: Ashis Sau, ashis.sau@tcs.com

**Additional Contacts**: Suraj Mahato, surajkumar.mahato@tcs.com

## 1 Use Case Summary Table

| Item | Details |
|---|---|
| Category | Climate Change/Natural Disaster |
| Problem Addressed | In case of massive disaster (like flood) over a vast area, the terrestrial cellular connectivity may get destroyed. This jeopardizes the ability of remote instantaneous visual tracking of the disaster site by the central control room. Even sending HD images or videos over Geostationary (GEO) satellite based communication channel leads to high latency and suffers from lack of bandwidth in practice. This leads to delayed comprehensive disaster response strategy.<br>In this context we propose a low-bandwidth consuming system based on the principles of semantic communication, coupled with GenAI based rendering to transfer 'semantically matching' live visualization of the remote disaster.<br>It may be noted that although Low Earth Orbit (LEO) satellites are promising for providing broad-band internet service for future, but back-haul on LEO has several operational, privacy and security issues due to their low-footprint and mobility. This also gives rise to many challenges due to the handover requirements and inter-satellite link congestion. So, GEO satellites will continue to be a reliable backhaul for critical utility services utilizing non-terrestrial network (NTN) [1]. |

(continued)

| Item | Details |
|---|---|
| Key Aspects of Solution | 1) Firstly, a common data set, consisting of regional images and location-based demographic information, to be created for transmitter side and receiver side containing regional<br>2) Then both the descriptor (at the Tx side) and visualizer (at the Rx side) LLMs are trained and fine tuned<br>3) Afterwards the descriptor module creates description of the visuals which is transmitted over the GEO Sat link.<br>4) The visualizer module consumes the receive description and generates an image/ video frames semantically matching to the source visuals. |
| Technology Keywords | Generative AI, Large-language models, Semantic communication |
| Data Availability | Private |
| Metadata (Type of Data) | Images, Location tag (optional, but desired) |
| Model Training and Fine-Tuning | LLaMA model (version 3 released on Apr'24, version 3.1 released on July'24, version 3.2 released on Sept'24) model is used for fine-tuning using custom dataset. |
| Testbeds or Pilot Deployments | The concept is yet to be fully deployed. The concept has been published in [2]. |

# 2 Use Case Description

## 2.1 Description

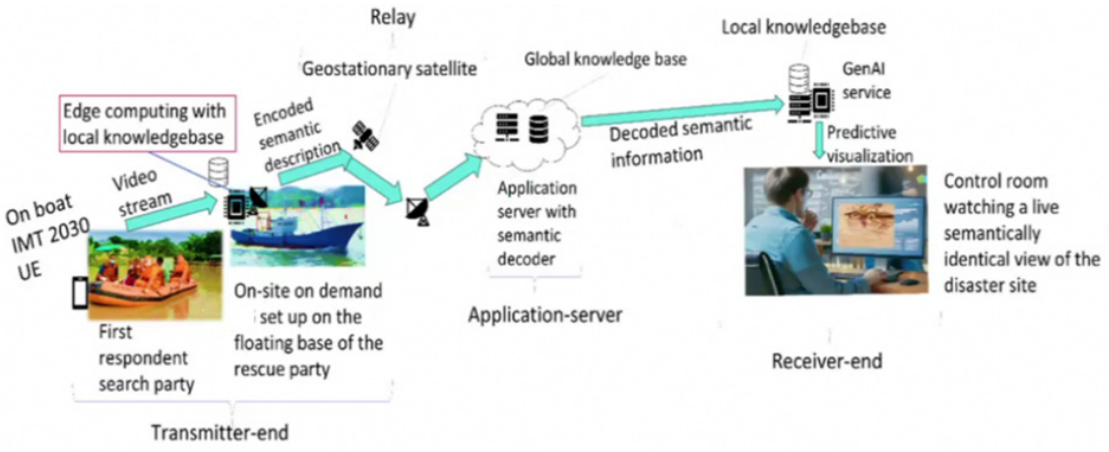
Let us consider a scenario of a massive disaster (like flood) over a vast area, the terrestrial cellular connectivity may get destroyed. This jeopardizes the ability of remote instantaneous visual tracking of the disaster site by the central control room. An alternative might be that a local first responded team transmits the visuals from the disaster sight using geostationary (GEO) satellite backhaul from the rescue boat. Although Low Earth Orbit (LEO) satellites are promising for providing broad-band internet service for future, but back-haul on LEO has several privacy and security issues due to their vicinity to earth and mobility. This also gives rise to many challenges due to the handover requirements and inter-satellite link congestion. GEO satellites continue to be a reliable backhaul for critical utility services utilizing non-terrestrial network (NTN) .

Due to the huge latency and lack of bandwidth over GEO based backhaul, sharing high-definition video and images in real-time may not be practically possible. As an alternative, we can conceive the Semantic Communication based futuristic disaster tracking system as in Fig. 1. In this case a private network is set up with non-terrestrial backbone. The network access point is installed on the floating large boat acting as the base unit of the recue party. The base is equipped with an antenna connected to the control unit through GEO. The rescue party roams around in a small boat equipped with user equipment (UE) which captures the live visuals and transmits to the base unit edge computing service. The semantic encoder residing in the edge computer extracts the relevant scene understanding and creates a semantic description against each visual. The encoded semantic information demands much less bandwidth compared to transmission of the full video or the image and can be transmitted over the GEO backhaul much

quicker with reliability. The decoded semantic information may either be directly consumed by the control station, or it can be fed to a GenAI server with local knowledge base to predictively recreate a scene corresponding to the received semantics.

If low semantic noise is ensured, then such a system can provide efficient remote visualization of disaster sites.

Thus, faster accurate response planning can be ensured.




This solution will benefit better response to massive disaster thereby potentially improving efficiency in well-being of life on land. Through better possibility of reaching out to the disaster response agency for agile actions, this solution leads to reduced impact of aftermath of disasters.

**Use Case Status:** Initial PoC Phase

## 2.2 Benefits of use case

This solution will benefit better response to massive disaster thereby potentially improving efficiency in well-being of life on land.

Through better possibility of reaching out to the disaster response agency for agile actions, this solution leads to reduced impact of aftermath of disasters.

## 2.3 Future Work

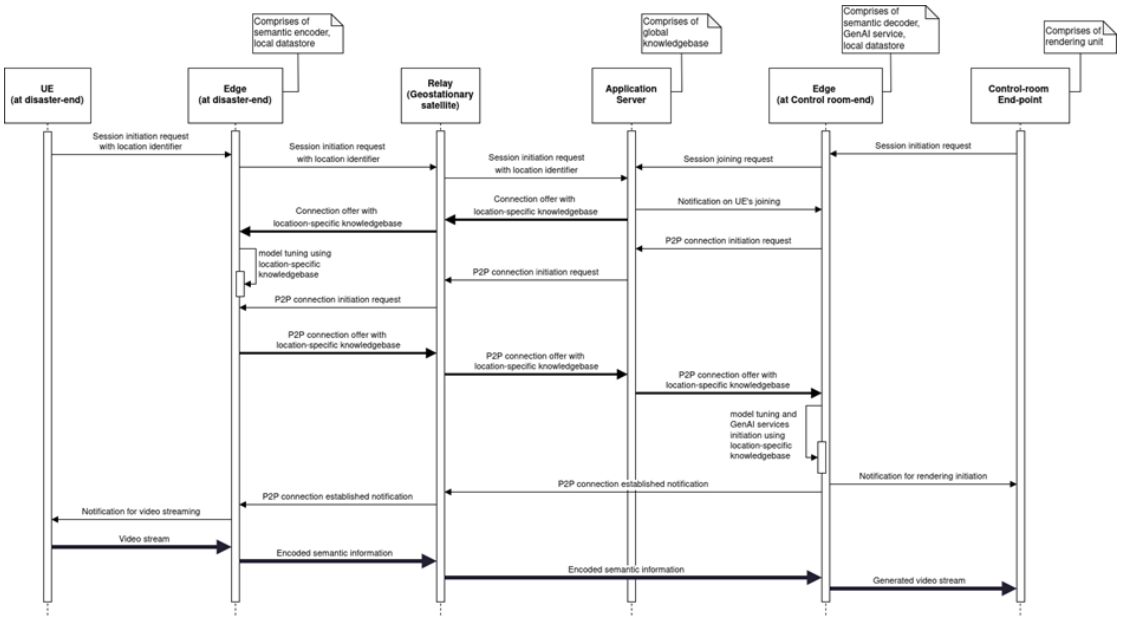
- Our initial finding shows that the existing datasets are inefficient for the purpose, especially considering countries like India and other developing and under-developed nations. So we need to develop a data set.
- We plan to use LlAMA as this is an open system and tune this with the data sets for the purpose. [3]
- Edge computing based end-to-end implementation

# 3 Use Case Requirements

- **REQ-01**: It is required for the UE to have on demand access to an MEC, if it is resource constrained.
- **REQ-02**: It is required for the models used in transmitter-side inference, receiver side decoding and GenAI to be seeded with similar data and training.

- **REQ-03**: It is required for the semantic representation at the transmitted to be encoded at the application layer with forward-error-correction (FEC) to ensure *in situ* recovery of lost packets and to meet latency requirements.
- **REQ-04**: It is critical for the network to have sufficient bandwidth to allow exchange of knowledgebases.
- **REQ-005**: It is recommended that the GenAI model may be tuned/ re-trained in runtime to enable narrowing down the search space for specific output generation utilizing specific information from the transmitter (e.g. location, time of the year, etc.).

## 4 Sequence Diagram




## 5 References

[1] P. Yue et al., "Low Earth Orbit Satellite Security and Reliability: Issues, Solutions, and the Road Ahead," in IEEE Communications Surveys & Tutorials, vol. 25, no. 3, pp. 1604-1652, thirdquarter 2023.

[2] A. Bhattacharyya, A. Sau and S. Mahato, "Semantic Communication with Networked AI in 6G: Application Scenarios, Insights from Practical Experiments, Standardization Aspects," 2025 17th International Conference on COMmunication Systems and NETworks (COMSNETS), Bengaluru, India, 2025, pp. 1264-1268, doi: 10.1109/COMSNETS63942.2025.10885668.

[3] Meta, "Llama 3.2," *Meta Llama*, 2022. https://www.llama.com/

# Use Case 15: AI Enables New Energy Meteorological Power Forecast

**Organization:** China Huadian Corporation LTD. & Huawei Technologies Co., Ltd.

**Country**: China

**Contact Person(S)**:

**Primary**: Fang Peng, fang-peng@chd.com.cn

**Secondary**: Hao Wen, wenhao23@huawei.com

## 1 Use Case Summary Table

| Item | Details |
|---|---|
| Category | Climate Change/Natural Disaster |
| Problem Addressed | Under the overarching goal of global sustainable development, the installed capacity of renewable energy has been experiencing rapid expansion worldwide. Both regulatory assessments and electricity market operations now require renewable energy operators to achieve significantly enhanced forecasting accuracy and operational efficiency. Nevertheless, the inherent characteristics of renewable sources, including the randomness, volatility, and intermittency of wind and solar power generation, pose substantial challenges to power grid dispatching systems. These challenges simultaneously create pressing demands for the advancement of traditional numerical weather prediction models to better accommodate the integration of renewable energy into modern power systems.<br>Traditional numerical weather forecasting predicts changes in wind speed and irradiance. This model, based on atmospheric dynamics, is complex and has not seen new breakthroughs in the past decade. It takes 6 to 10 hours to complete a prediction, which is time-consuming and has poor timeliness. The limited improvement in mechanism models and poor computational timeliness both affect the final prediction accuracy. |
| Key Aspects of Solution | 1) AI Large Model for meteorological forecasting.<br>2) Cloud-edge collaboration architecture (Group meteorological model, regional micro meteorological engine, site edge power prediction service). |
| Technology Keywords | AI Large Model, Cloud-Edge Collaboration, Meteorological Forecasting, Renewable Energy, Power Grid Optimization, Electricity Trading |
| Data Availability | Private |

(continued)

| Item | Details |
|---|---|
| Metadata (Type of Data) | 1) Meteorological data including time-series measurements (e.g., wind speed, solar irradiance, temperature, humidity) and geospatial data (e.g., longitude-latitude coordinates).<br>2) Data formats include NetCDF for large-scale meteorological datasets and CSV for site-specific records.<br><br>Metadata includes attributes such as timestamp, geographic coordinates (latitude, longitude), sensor type, measurement unit, and data collection frequency (e.g., 15-minute intervals). |
| Model Training and Fine-Tuning | The AI Large Model, based on Huawei's Pangu model [1], is pre-trained on global meteorological datasets and fine-tuned using transfer learning on proprietary datasets from China Huadian and Beijing Jiutian. Fine-tuning involves supervised learning with historical meteorological and power generation data to optimize predictions for wind speed, solar irradiance, and power output. No Retrieval-Augmented Generation (RAG) or external knowledge base is used; the model relies on end-to-end training with domain-specific data. |
| Testbeds or Pilot Deployments | Pilot verification at Huadian Jiangsu pilot station, reducing assessment fines in the first half of 2024. Based on Huawei high-performance AI computing ability, including training and inference ability, Our solution provides different deployment policies in different scenarios. At Huadian, the AI model runs on a dedicated computing cards and does not require supercomputer clusters. The entire deployment is efficient and highly scalable. |

# 2 Use Case Description

## 2.1 Description

**Introduction**: Amid the global push for sustainable development, the installed capacity of renewable energy continues to grow rapidly. However, the inherent unpredictability, variability, and intermittency of wind and solar power pose major challenges to power grid dispatching. Meanwhile, both regulatory assessments and electricity market trading demand higher forecasting precision and efficiency from renewable energy operators—requirements that traditional numerical weather prediction models struggle to meet. This use case aims to enhance the accuracy and efficiency of new energy meteorological power forecasts to support grid stability, optimize electricity trading, and reduce assessment penalties.

**Solution Overview**: China Huadian, in collaboration with Huawei, developed a new energy meteorological power prediction solution using a cloud-edge collaboration three-layer architecture:

- Group meteorological model: Processes large-scale historical and real-time meteorological data.
- Regional micro meteorological engine: Refines forecasts for specific regions.
- Site edge power prediction service system: Delivers localized, real-time predictions.
- The solution utilizes AI Large Model technology (e.g., Huawei's Pangu model), which outperforms traditional numerical models by completing global high-resolution forecasts in seconds, nearly 10,000 times faster. This approach achieves forecast accuracies of

97.24% (15-minute), 91.72% (4-hour), and over 90% (24-hour), significantly improving timeliness and precision.

The expected impact and benefits of the solution:

- Improved Forecast Accuracy: Enhances grid dispatching reliability, optimizes electricity trading, and reduces penalties
- Cost Savings: Pilot data from Huadian Jiangsu shows assessment penalties reductions of 27% and 15% in the first half of 2024.
- Sustainability: Supports efficient integration of renewable energy, contributing to global sustainable development and clean energy goals.
- Market Efficiency: Enables power companies to make informed trading decisions, improving economic outcomes in electricity markets.
- Scalability: The low-cost AI model runs on dedicated computing cards, eliminating the need for supercomputer clusters.

The AI Large Model-based forecasting solution enhances the reliability and efficiency of renewable energy integration into the power grid, supporting affordable and clean energy. By improving forecast accuracy, it reduces operational costs and penalties, making renewable energy more economically viable. It also contributes to climate action by enabling higher renewable energy adoption, reducing fossil fuel reliance, and supporting global carbon neutrality goals. The rapid computation and low-cost infrastructure democratize access to advanced forecasting, promoting sustainable energy practices worldwide.

**Use Case Status:** The use case is part of a larger research project.

Partners:

- **China Huadian Corporation LTD.:** Provides critical data from stations, enabling precise forecasts as we saw in earlier results. Provides computing platform for model training and selects two pilot sites to verify the model prediction effect.
- **Huawei Technologies Co., Ltd.:** Provides unified digital infrastructure and develops and deploys the AI system.
- **Beijing Jiutian Meteorological Technology Co., Ltd.:** Provides meteorological data and supports modeling from a meteorological perspective, ensuring accurate and relevant data inputs for the AI Large Model.

## 2.2 Benefits of use case

The AI Large Model-based forecasting solution enhances the reliability and efficiency of renewable energy integration into the power grid, supporting affordable and clean energy. By improving forecast accuracy, it reduces operational costs and penalties, making renewable energy more economically viable. It also contributes to climate action by enabling higher renewable energy adoption, reducing fossil fuel reliance, and supporting global carbon neutrality goals. The rapid computation and low-cost infrastructure democratize access to advanced forecasting, promoting sustainable energy practices worldwide.

## 2.3 Future Work

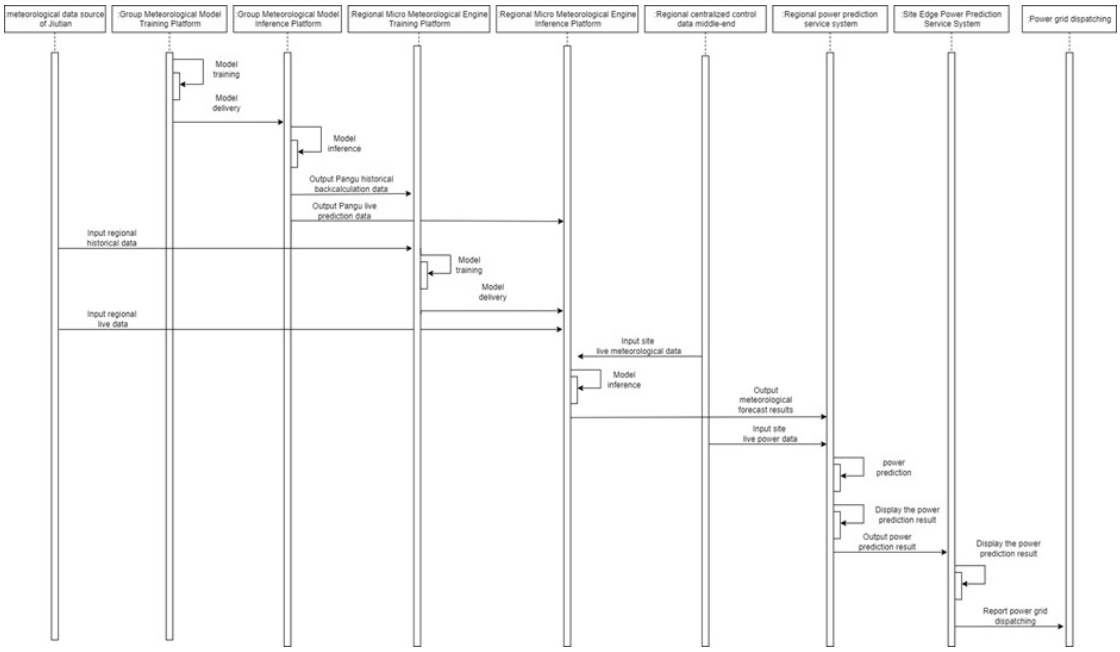
The future work of AI enables new energy meteorological power forecast includes:

- Expand prediction capabilities to include precipitation and temperature forecasts, impacting power grid load and trading decisions.
- Extend forecast duration from medium-short-term (within 10 days) to cross-month and cross-quarter predictions to guide long-term power trading and grid planning.
- Develop standards for AI-based meteorological forecasting in the energy sector.
- Develop a standardized knowledge base for AI-based meteorological forecasting in collaboration with industry partners and regulatory bodies. This knowledge base will include standardized data formats metadata schemas, and forecasting protocols to ensure interoperability and scalability across renewable energy operators. The goal is to establish an industry benchmark for AI-driven meteorological forecasting, promoting adoption and consistency in the energy sector.
- Enhanced datasets for precipitation, temperature, and market-related data from partners.
- Develop advanced model for long-term forecasting.
- Deepen partnerships with meteorological firms to enhance data quality and modeling support.

# 3 Use Case Requirements

- **REQ-01**: It is required to have access to large-scale historical and real-time meteorological data from partners like Beijing Jiutian Meteorological Technology Co., Ltd. for training and fine-tuning the AI Large Model.
- **REQ-02**: It is required to have cloud-edge collaboration infrastructure to support the three-layer architecture (group, regional, and site-level processing).
- **REQ-03**: It is required to have high-performance AI computing ability for training the meteorological model efficiently.
- **REQ-04**: It is required to have integration with power station systems and electricity trading platforms to deliver real-time forecast data for grid dispatching, reporting, and trading.
- **REQ-05**: It is required to comply with regional energy regulations (e.g., Jiangsu Province's forecasting standards) and electricity market requirements to avoid penalties and optimize trading.

## 4 Sequence Diagram




## 5 References

[1] Bi, K., Xie, L., Zhang, H. et al. Accurate medium-range global weather forecasting with 3D neural networks. Nature 619, 533–538 (2023). https://doi.org/10.1038/s41586-023-06185-3

# Use Case 16: AI-Driven Early Warning System for Oceanic Dead Zones

**Organization:** Vellore Institute of Technology, Chennai Campus

**Country:** India

**Contact Person(s):**

Dr. J Deepika Roselind, deepikarooselind.j@vit.ac.in

## 1 Use Case Summary Table

| Item | Details |
|---|---|
| Category | Climate Change/Natural Disaster |
| Problem Addressed | Increasing prevalence of oceanic dead zones due to climate change and nutrient runoff, causing marine biodiversity loss and impacting coastal economies. |
| Key Aspects of Solution | AI(Artificial Intelligence)-based early warning system using satellite and sonar data.<br>Deep learning models to predict low-oxygen zones 2–4 weeks in advance.<br>Real-time alert system for proactive mitigation. |
| Technology Keywords | AI, machine learning, oceanography, satellite data, sonar, hypoxia prediction, climate analytics, early warning systems |
| Data Availability | Public and Private<br>National Aeronautics and Space Administration(NASA) OceanColor: [1]<br>National Oceanic and Atmospheric Administration National Centers for Environmental Information(NOAA NCEI): [2]<br>Copernicus Marine Environment Monitoring Service(CMEMS): [3] |
| Metadata (Type of Data) | Images (satellite imagery), environmental sensor data (temperature, dissolved oxygen, chlorophyll), spatiotemporal marine measurements |
| Model Training and Fine-Tuning | Deep learning (Convolutional Neural Networks(CNNs)), multivariate regression and anomaly detection models trained on historical hypoxia data from global marine datasets |

(continued)

| Item | Details |
|---|---|
| Testbeds or Pilot Deployments | Government agencies such as National Institute of Oceanography (NIO) and Earth System Science Organization (ESSO) – Indian National Centre for Ocean Information Services (INCOIS) |

## 2 Use Case Description

### 2.1 Description

Oceanic "dead zones" are areas in the ocean where oxygen levels drop below the level required to support marine organisms. These hypoxic areas are mainly caused by ocean warming from climate change, nitrogen and phosphorus-rich agricultural runoff, and uncontrolled urban pollution. Consequently, dead zones are growing in both extent and occurrence, presenting significant risks to worldwide biodiversity, depleting essential fisheries, and undermining coastal economies—especially in sensitive regions like the Gulf of Mexico, the Baltic Sea, and the Arabian Sea. In response to this escalating crisis, the aim of this use case is to create and implement an AI-driven early warning system that can forecast the emergence of oceanic dead zones 2 to 4 weeks ahead of time, enabling stakeholders to take proactive measures before permanent harm takes place. This system is designed to strengthen marine conservation initiatives, improve climate resilience, and protect food security by utilizing advanced, real-time ocean monitoring integrated with automated alert systems.

The proposed solution combines multi-source environmental data gathered from satellite remote sensing platforms (e.g., Moderate Resolution Imaging Spectroradiometer on the Aqua(MODIS Aqua), Sentinel-3), fixed buoy networks with sensors for dissolved oxygen, temperature, salinity, and chlorophyll, autonomous underwater vehicles (AUVs) like SeaGliders, sonar-based ocean profilers,Conductivity-Temperature-Depth (CTD) profilers, and Argo floats. These sensors together monitor essential environmental indicators such as sea surface temperature (SST), chlorophyll levels (indicative of phytoplankton biomass), dissolved oxygen, salinity variations, nutrient levels (nitrate and phosphate), and water density layering.

Sophisticated AI models—especially CNNs for identifying spatial patterns and long short-term memory (LSTM) networks for learning temporal sequences—are employed to examine these extensive spatiotemporal datasets and identify early indicators of hypoxia, facilitating high-confidence predictions of potential risk areas. The predictive results are shared via a multilingual decision-support dashboard, enhanced by real-time email alerts and Short Message Service(SMS) early warning systems, allowing policymakers, fisheries management bodies, and conservation groups to quickly adopt interventions like nutrient management strategies, artificial oxygenation initiatives, or temporary fishing restrictions. This all-encompassing strategy aims to greatly minimize the decline of marine species, guide sustainable fisheries policies, bolster environmental resilience against climate change.

The expected impacts of this system include:

- Reduction of marine life loss and preservation of biodiversity.
- Improved Policy Formulation for sustainable and adaptive fisheries management.

- Strengthened environmental resilience to climate-driven oceanic changes.

Additionally, the system promotes open science and reproducibility by committing to share curated datasets through platforms such as Institute of Electrical and Electronics Engineers(IEEE) DataPort for further research and cross-border collaboration.

**Use Case Status:** The system is currently at the prototype stage with data ingestion pipelines, AI models, and validation frameworks in place. Pilot deployments are planned for hypoxia-prone regions like the Gulf of Mexico, Bay of Bengal, and Arabian Sea.

**Partners:** National Oceanic and Atmospheric Administration(NOAA), European Space Agency(ESA), Global Ocean Oxygen Database and Atlas(GO2DAT), National Institute of Ocean Technology(NIOT), Earth System Science Organization & National Centre for Ocean Information Services(ESSO & NCOIS)

## 2.2 Benefits of use case

The system supports efforts to reduce marine pollution, particularly from land-based sources, by detecting early signs of hypoxia in marine environments. This enables timely interventions that can prevent widespread fish mortality and protect sensitive habitats. By integrating real-time environmental data with predictive AI capabilities, the platform enables targeted regulatory actions and promotes evidence-based marine policy enforcement. These functions directly contribute to the sustainable management and conservation of marine ecosystems, helping to restore degraded environments and maintain biodiversity.

In addition, the AI solution enhances fisheries management in hypoxia-prone regions. Through advanced forecasting and analysis, it supports the regulation of harvesting activities, helping to prevent overfishing and ensure the long-term viability of fish stocks. This contributes to the ecological and economic sustainability of marine resources.

Ocean dead zones, exacerbated by climate change-related warming, pose growing risks to marine ecosystems and food systems. The forecasting capabilities of this system provide critical tools for adaptive marine governance. By enabling localized responses to oxygen depletion and other climate-driven hazards, the platform supports climate resilience in vulnerable coastal regions.

The system also strengthens early warning mechanisms for environmental degradation, improving national and regional readiness to respond to climate-related challenges. Its integration into planning processes can help governments incorporate real-time ocean health data into national policies and adaptation strategies, enhancing their overall climate response capabilities.

By protecting coastal fisheries from collapse due to oxygen depletion, the platform safeguards a vital source of food and income for coastal communities. This directly contributes to maintaining stable marine-based food supplies, which are essential for ensuring food security, especially in low-income and climate-vulnerable areas.

Moreover, the system supports the productivity and sustainability of small-scale fisheries. By providing actionable data and timely alerts, it enables fishers to adapt their practices, avoid ecological damage, and maintain their livelihoods. This targeted support reinforces sustainable food production and strengthens the economic resilience of marine-dependent communities.

## 2.3 Future Work

- Enhanced Geographic Coverage: Future deployments aim to extend the system's coverage to additional hypoxia-prone regions, such as the Black Sea, South China Sea, and Indian coastal waters. This requires access to localized datasets and collaboration with regional oceanographic institutes.
- Integration of Autonomous Monitoring Technologies: Incorporation of autonomous drones and underwater robotic platforms equipped with real-time oxygen and nutrient sensors will improve spatial and temporal granularity of data. These platforms can fill gaps in satellite and buoy coverage, particularly in remote or under-monitored marine regions.
- Real-Time Intervention Feedback Mechanism: Developing a two-way feedback interface where field-level marine workers and local agencies can directly report outcomes and anomalies back into the AI system will enhance model responsiveness and regional specificity.
- Multilingual and Inclusive Alert System: Expanding the system to support multilingual alerts and low-tech communication channels (e.g., SMS) will ensure that small-scale fisherfolk and community stakeholders are informed and can act quickly.
- Model Explainability and Transparency: Building tools for explainable AI (XAI) to help users and decision-makers understand why certain regions are flagged as high risk will increase trust, adoption, and model auditing capacity.

# 3 Use Case Requirements

The successful deployment of the AI-Driven Early Warning System for Oceanic Dead Zones relies on a combination of technical infrastructure, environmental data streams, institutional partnerships, scalable Application Programming Interfaces(APIs), and responsible policy frameworks. The following essential requirements are identified:

**REQ-01**: Real-Time Environmental Data Acquisition

Real-time satellite imagery (sea surface temperature, ocean colour, chlorophyll levels) and in-situ oceanographic sensor feeds (dissolved oxygen, salinity, pH, nutrient levels) are required from organizations such as NASA, NOAA, ESA, NIOT, and ESSO-NIOT–NCOIS. Additional data will be collected using AUVs like SeaGliders, underwater drones, CTD profilers, MODIS Aqua satellite feeds and Argo floats to enhance spatial and temporal resolution.

**REQ-02**: High-Performance Computational Infrastructure

Cloud or on-premises systems equipped with Graphics Processing Units(GPUs) or Tensor Processing Units(TPUs) are required to facilitate training and inference of deep learning models on large-scale, multi-source ocean datasets. The system architecture must support scaling for high-volume geospatial data.

**REQ-03**: Multi-Source Data Integration & Preprocessing Pipeline

The system must support ingestion and integration of heterogeneous data formats, including Network Common Data Form(netCDF)(satellite data)), Comma-Separated Values(CSV) (sonar data)), and real-time APIs (NASA EarthData, NOAA NCEI File Transfer Protocol(FTP), CMEMS RESTful services). Automated pipelines will normalize, align, and fuse datasets dynamically.

**REQ-04**: AI Model Development, Retraining, and Continuous Learning

Deep learning pipelines incorporating CNNs, LSTMs, and ensemble architectures must be modular and support continual retraining based on incoming new environmental data and real-world validation feedback. Models should adapt to regional oceanographic changes to maintain predictive accuracy.

**REQ-05**: Decision-Support Dashboard and Stakeholder Alert System

A dynamic geospatial dashboard will visualize hypoxia risk forecasts. Alerts must be disseminated to stakeholders—marine agencies, fisheries departments, and coastal governance bodies—via SMS notifications, email alerts and mobile dashboards in multiple languages to ensure widespread, timely action.

**REQ-06**: Field Validation, Sensor Calibration, and Pilot Deployments

Collaborations with regional marine research institutes are required for physical field validation, sensor calibration (buoys, AUVs, CTDs), and pilot deployments in hypoxia-prone regions such as the Gulf of Mexico, Bay of Bengal, and Arabian Sea.

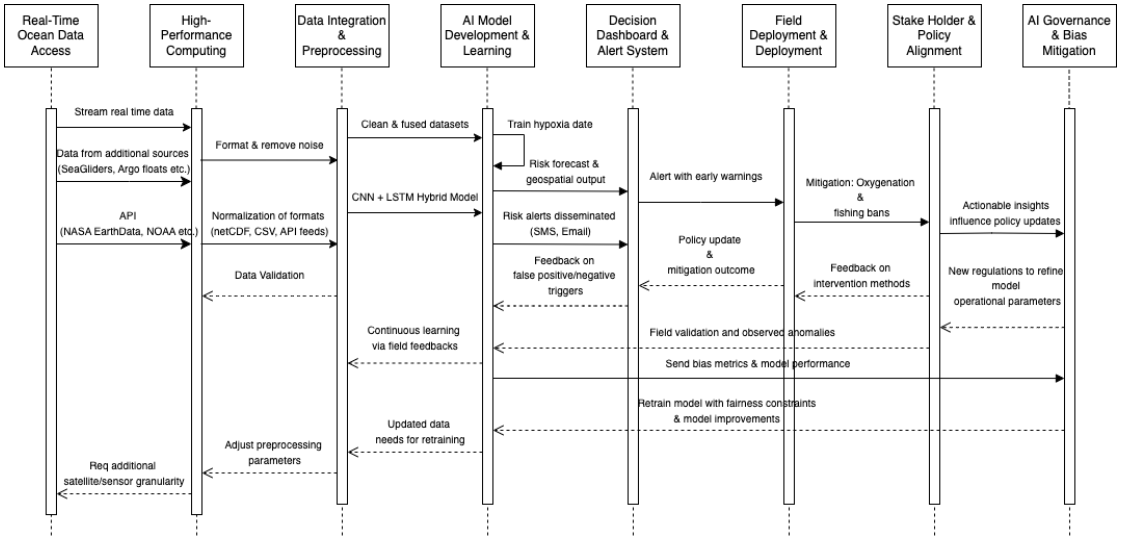
**REQ-07**: Policy Integration, Institutional Cooperation and Governance

Institutional cooperation is essential to align AI model outputs with marine conservation policies, fisheries regulations, pollution control strategies. Outputs will also support national ocean service frameworks and open governance portals.

**REQ-08**: Ethical AI Governance and Open Data Publication

The system must ensure responsible AI practices, including fairness audits, transparency reporting, and inclusivity for underrepresented ocean regions in model training. Newly generated datasets will be openly shared via platforms like IEEE DataPort and global ocean research repositories to foster replicability and community engagement.

## 4 Sequence Diagram



## 5 References

[1] NASA. (n.d.). NASA OceanColor Web. National Aeronautics and Space Administration. Retrieved from https://oceancolor.gsfc.nasa.gov

[2] National Oceanic and Atmospheric Administration. (n.d.). National Centers for Environmental Information (NCEI). Retrieved from https://www.ncei.noaa.gov

[3] Copernicus Marine Environment Monitoring Service. (n.d.). Ocean monitoring and forecasting. European Union. Retrieved from https://marine.copernicus.eu

[4] Global Ocean Oxygen Database and Atlas. (n.d.). GO2DAT: A repository of ocean oxygen data. Retrieved from https://www.go2dat.org

[5] NOAA. (2023, June). NOAA forecasts approximately average-sized Gulf of Mexico dead zone this summer. National Oceanic and Atmospheric Administration Newsroom. Retrieved from https://www.noaa.gov/news-release/noaa-forecasts-approximately-average-sized-gulf-of-mexico-dead-zone-this-summer

[6] Li, Y., Robinson, S. V. J., Nguyen, L. H., & Liu, J. (2023). Satellite prediction of coastal hypoxia in the northern Gulf of Mexico. Remote Sensing of Environment, 284, 113346. https://doi.org/10.1016/j.rse.2022.113346

[7] Nieves, V., Ruescas, A., & Sauzède, R. (2024). AI for marine, ocean and climate change monitoring. Remote Sensing, 16(1), 15. https://doi.org/10.3390/rs16010015

# Use Case 17: AI-Powered RF Signal Locator for Advanced Search and Rescue Operations

**Organization:** Mbeya University of Science and Technology (MUST)

**Country:** Tanzania

**Contact Person(s):**

Faisali Malukula Gwaya, faisali4gwaya@gmail.com

## 1 Use Case Summary Table

| Item | Details |
|---|---|
| Category | Climate change and natural disaster |
| Problem Addressed | Challenges in accurately locating individuals during SAR missions due to insufficient real-time location information. |
| Key Aspects of Solution | AI-based RF signal triangulation; Real-time data analysis and location prediction; Seamless integration with SAR operational protocols. |
| Technology Keywords | Artificial Intelligence, RF Localization, Machine Learning, Search and Rescue, Signal Processing |
| Data Availability | Private (planned data collection during pilot testing) |
| Metadata (Type of Data) | RF Signal Strength, GPS Coordinates, Timestamps, Sensor Metadata |
| Model Training and Fine-Tuning | Supervised learning models trained on synthetic and real RF datasets, with transfer learning for terrain adaptation. |
| Testbeds or Pilot Deployments | Pilot tests in collaboration with SAR teams under controlled environments. |
| Code Repositories | Planned: Repository access post-validation. |

## 2 Use Case Description

### 2.1 Description

This use case proposes an AI-enhanced solution to efficiently locate individuals during search and rescue missions by triangulating RF signals emitted from mobile devices or beacons. Traditional SAR relies heavily on visual and audio clues, which can delay rescues. AI models analyzing RF signal patterns will significantly boost response accuracy and timing.

The machine learning models will use environmental and signal data to predict the missing persons' location, improving survival chances and operational efficiency.

**Use Case Status:** Planning Phase

## 2.2 Benefits of use case

Sustainable Cities and Communities: Enhancing emergency response infrastructure. Good Health and Well-Being: Facilitating timely rescues and health outcomes during emergencies.

- Project Partners Mbeya University of Science and Technology (MUST): Research and development of AI localization technologies.
- Telecom Company: Provision of RF data and communication infrastructure support.
- Rescue Team: Field testing, feedback, and operational deployment of the solution.

## 2.3 Future Work

- Define system specifications.
- Collaborate for data acquisition with SAR teams.
- Develop AI models and validate through field tests.
- Expand system for UAV-based aerial searches.

# 3 Use Case Requirements

- REQ-01: Reliable RF signal triangulation across terrains.
- REQ-02: Real-time predictions with minimal computational resources.
- REQ-03: Portability and SAR kit compatibility.
- REQ-04: Aerial and ground operational integration.

# 4 Sequence Diagram

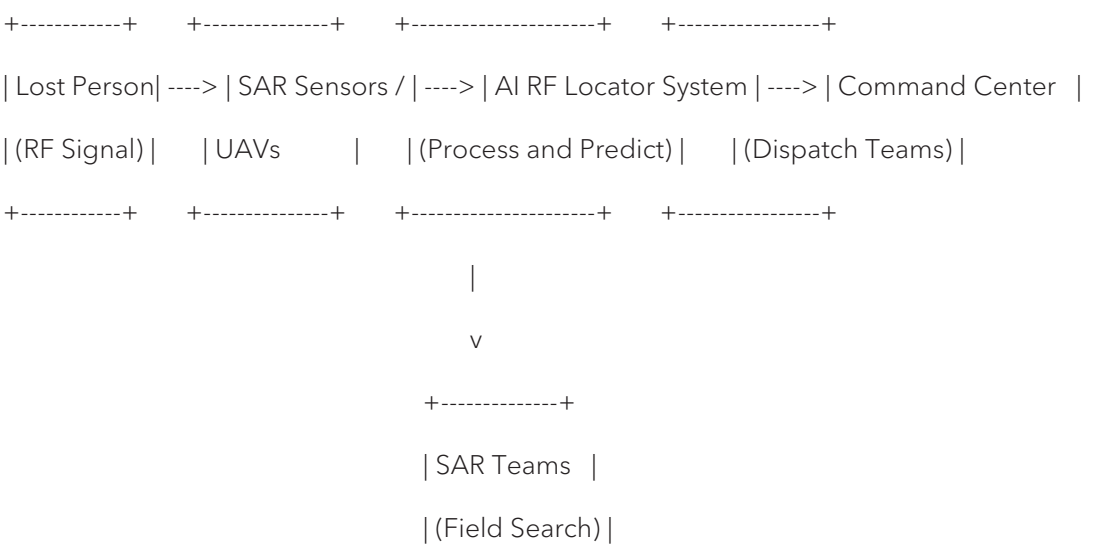
Actors:

- Lost Individual (RF Signal Source)
- SAR Teams (Ground and UAV)
- AI RF Locator System
- Command Center

Process Flow:

1. RF signal emitted by lost individual.
2. Signal detection by sensors or UAVs.
3. Signal data sent to AI RF Locator.
4. AI processes data and predicts location.
5. Location shared with Command Center.
6. SAR Teams dispatched based on AI guidance.

7. Search conducted and status updates sent.




## 5 References

[1] Rappaport, Theodore S. *Wireless Communications: Principles and Practice*. 2nd ed. Upper Saddle River, NJ: Prentice Hall, 2002.

[2] Bahl, Paramvir, and Venkata N. Padmanabhan. "RADAR: An In-Building RF-Based User Location and Tracking System." In *Proceedings of IEEE INFOCOM 2000*, 775–784. Tel Aviv, Israel: IEEE, 2000.

[3] International Telecommunication Union (ITU). *ITU-R Recommendation P.1238-9: Propagation Data and Prediction Methods for the Planning of Indoor Radiocommunication Systems and Radio Local Area Networks in the Frequency Range 300 MHz to 100 GHz*. Geneva: ITU, 2017.

[4] Roos, Teemu, Perttu Myllymäki, Hannu Tirri, Petri Misikangas, and Juha Sievänen. "A Probabilistic Approach to WLAN User Location Estimation." *International Journal of Wireless Information Networks* 9, no. 3 (2002): 155–164.

[5] International Federation of Red Cross and Red Crescent Societies (IFRC). *Search and Rescue Operations Guidelines*. Geneva: IFRC, 2018.

[6] Institute of Electrical and Electronics Engineers (IEEE). *IEEE Standard for Information Technology–Telecommunications and Information Exchange between Systems–Local and Metropolitan Area Networks–Specific Requirements–Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications*. IEEE Std 802.11-2020. New York: IEEE, 2020.

[7] Liu, Hui, Houshang Darabi, Pat Banerjee, and Jing Liu. "Survey of Wireless Indoor Positioning Techniques and Systems." *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)* 37, no. 6 (2007): 1067–1080.

# Use Case 18: Cooling the Concrete Jungle: An AI-Enabled Urban Heat Island Mitigation Advisor for Sustainable Development.

**Organization**: Christ University, Bangalore, India

**Country**: India

**Contact Person(s)**:

Primary: Satyam Kumar, satyamkumar0905@outlook.com

Secondary: Priyansha Upadhyay, priyansha.upadhyay1304@gmail.com

## 1 Use Case Summary Table

| Item | Details |
|---|---|
| Category | Climate change and natural disaster |
| Problem Addressed | Urban Heat Island (UHI) effect causing dangerous temperature increases in cities, leading to health emergencies disproportionately impacting vulnerable populations |
| Key Aspects of Solution | – Multi-source data integration (satellite imagery, IoT sensors, socioeconomic indicators)<br>– Advanced AI modeling (GANs, spatiotemporal models, explainable AI)<br>– Real-time hotspot prediction<br>– Interactive scenario modeling<br>– Policy and cost-benefit analysis<br>– Equity-prioritized intervention |
| Technology Keywords | Artificial Intelligence, Machine Learning, GAN, Spatiotemporal modeling, Explainable AI, IoT sensors, Satellite imagery, Real-time analytics, Smart city infrastructure |
| Data Availability | Mixed (Public weather data, satellite imagery; Private IoT sensor networks) |
| Metadata (Type of Data) | Geospatial, Temporal, Environmental, Demographic, Socioeconomic, Infrastructure |
| Model Training and Fine-Tuning | GANs for thermal mapping, spatiotemporal prediction models, ensemble and explainable AI for transparent support |
| Testbeds or Pilot Deployments | In development, MVP phase currently underway |
| Code repositories | GitHub: ITU_heat_island_ai_4_good |

# 2 Use Case Description

## 2.1 Description

The Urban Heat Island Mitigation Advisor addresses the growing threat of urban heat, which in 2024 alone led to 40,000+ suspected heat stroke cases in India. Urban areas are becoming thermal danger zones due to concrete infrastructure, climate change, and poor mitigation planning. UHI poses a severe public health challenge, especially for the elderly, children, low-income groups, and outdoor workers.

This AI-based system integrates diverse datasets—satellite images, IoT sensor feeds, socioeconomic indicators – and uses state-of-the-art AI (GANs, spatiotemporal prediction, explainable AI) to analyze and visualize real-time and forecasted UHI trends. It supports decision-makers in deploying timely interventions: green corridors, reflective materials, emergency resource allocation, and smart infrastructure.

**Use Case Status**: Pilot, part of a larger product development

**Category**: Climate resilience, Urban planning, Environmental monitoring, Public health

**Partners**: Internal development at Christ University with future collaboration opportunities

Use Case Status: Pilot, part of a larger product development

Partners: Internal development at Christ University with future collaboration opportunities

## 2.2 Benefits of use case

The Urban Heat Island Mitigation Advisor directly contributes to multiple Sustainable Development Goals through its integrated approach to addressing urban heat challenges:

- The system prioritizes public health by identifying vulnerable populations and targeting interventions to reduce heat-related illnesses and mortality. Incorporating health impact metrics and vulnerability assessments ensures that cooling measures protect those most at risk during extreme heat events.
- The tool reduces cooling-related energy demands by optimizing energy consumption by integrating building management systems and intelligent energy networks. This promotes energy efficiency and supports the transition to more sustainable urban energy systems.
- The solution exemplifies innovation in urban infrastructure by leveraging advanced AI technologies to enhance urban resilience. Its integration with existing smart city networks promotes adaptive, intelligent infrastructure systems.
- The advisor explicitly addresses inequality by identifying communities disproportionately affected by urban heat and prioritizing interventions in these areas. This equity-driven approach ensures that benefits reach marginalized populations facing the greatest climate-related risks.
- The tool promotes sustainable urban development by enhancing cities' resilience to heat stress. Scenario modeling and targeted interventions help create more livable, thermally comfortable urban environments for all residents.
- The solution addresses climate change adaptation by helping cities respond to increasing heat stress. Providing decision support for heat mitigation strategies strengthens urban resilience to climate-related hazards.

## 2.3 Future Work

Future work involves expanding the system into an Integrated Climate Resilience Platform to address multiple threats (floods, air quality, droughts). Enhancements include:

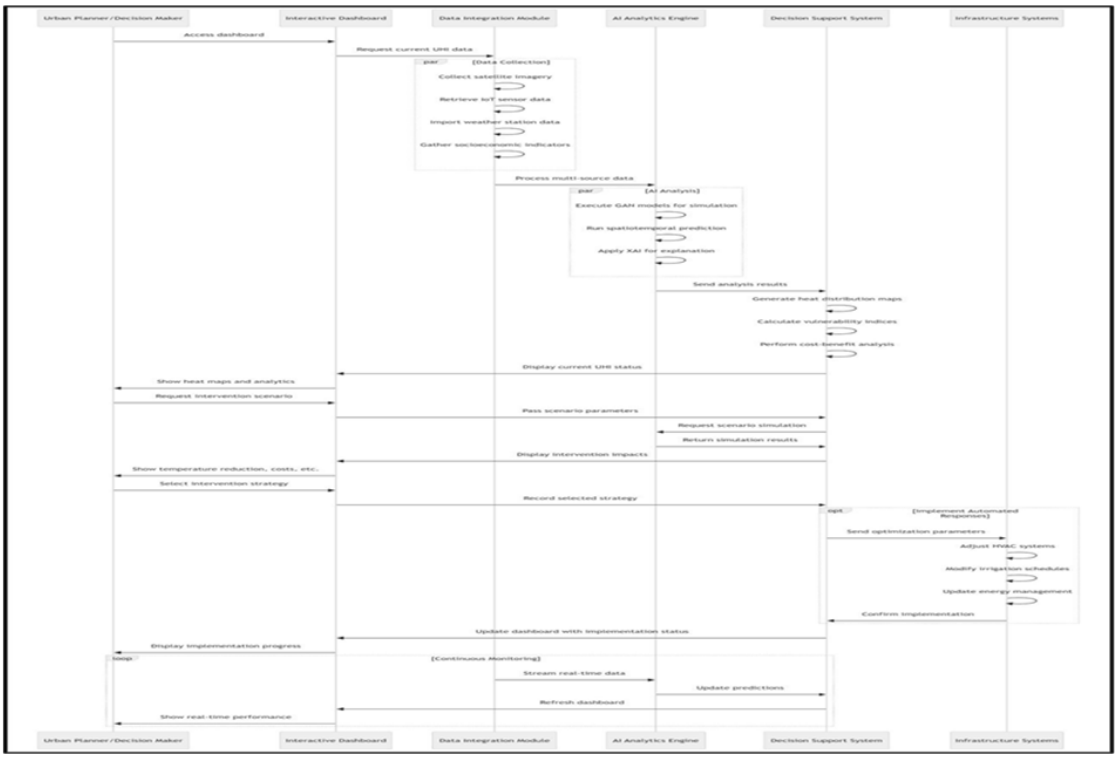
- Functionality: Multi-hazard prediction, ecosystem service valuation, policy optimization
- Components: Vulnerability mapping, AI-driven co-benefit assessment, community engagement modules

Resources Needed: New sensor integrations, improved models, expanded institutional partnerships

# 3 Use Case Requirements

- **REQ-01**: Integration of multi-source data (satellite, IoT, weather, socioeconomic)
- **REQ-02**: Use of advanced AI models: GANs, spatiotemporal models, explainable AI
- **REQ-03**: Real-time analytics for hotspot identification and prediction
- **REQ-04**: Interactive decision support and scenario simulation
- **REQ-05**: Cost-benefit and policy impact analysis
- **REQ-06**: Integration with smart infrastructure systems
- **REQ-07**: Equity-based prioritization tools
- **REQ-08**: Visualization of current and future UHI distributions
- **REQ-09**: Simulation engine for mitigation strategy evaluation
- **REQ-10**: Targeted planning tools (green corridors, cool roofs, etc.)

# 4 Sequence Diagram and dashboard



## 5 References

[1] Research on Urban Heat-Island Effect. Available online: https://doi.org/10.1016/j.proeng.2016.10.002

# Use Case 19: Spatio-Temporal Microclimate Prediction using Deep Learning

**Organization**: Sri Sivasubramaniya Nadar College of Engineering

**Country**: India

**Contact Person(s)**:

Syed Azim – syed2110753@ssn.edu.in

Srikanth S – srikanth2110893@ssn.edu.in

Prasanna Karthikeyan – prasanna2110778@ssn.edu.in

Kaythry P – kaythryp@ssn.edu.in

## 1 Use Case Summary Table

| Item | Details |
|---|---|
| Category | Urban Health & Comfort |
| Problem Addressed | Limited availability of localized microclimate data in urban areas due to sparse weather stations and high cost of dense sensor networks. |
| Key Aspects of Solution | – Vision Transformer model to predict microclimate parameters<br>– Combines satellite/street images with weather station data<br>– High-resolution weather maps without dense sensor deployment |
| Technology Keywords | Vision Transformer (ViT), Microclimate Prediction, Satellite Imagery, Spatio-Temporal Modeling |
| Data Availability | Public: Satellite imagery from Mapbox Static Tiles API<br>Private: On-ground weather data & street-level images |
| Metadata (Type of Data) | Visual: Satellite and street-level imagery<br>Numerical: Time-series weather data (temperature, humidity, wind speed, irradiance) |
| Model Training and Fine-Tuning | Vision Transformer (ViT) trained with multi-source visual and numeric data to capture spatial and temporal patterns |
| Testbeds or Pilot Deployments | ScienceDirect Publication[7] |
| Code Repositories | GitHub: Microclimate Prediction[8] |

# 2 Use Case Description

## 2.1 Description

Urban environments experience significant microclimate variability caused by buildings, vegetation, and human activity. Traditional weather systems fail to capture these localized effects, creating challenges for urban comfort, heat risk mitigation, and sustainable design. This use case aims to develop a scalable AI solution using Vision Transformers to predict microclimate parameters—temperature, humidity, wind speed, and solar irradiance, at high spatial and temporal resolution.

The model is trained on:

- Satellite imagery from Mapbox API
- Street-level images taken near weather stations
- Real-time weather station data from a pilot campus

Unlike conventional forecasting models, Vision Transformers (ViT) leverage self-attention mechanisms to learn spatial-temporal relationships from both image and numeric data. This enables precise mapping without dense sensor networks or computationally expensive simulations.

Expected Benefits:

- More informed urban planning and public space design
- Better outdoor thermal comfort for pedestrians and cyclists
- Improved climate resilience and resource-efficient agriculture
- Data-driven decision support for campus and urban development

**Use Case Status**: Operational prototype; academic deployment on a college campus

## 2.2 Benefits of use case

- Supports thermal comfort and reduces heat-related risks for vulnerable urban populations
- Promotes AI-driven climate monitoring innovations and resilient infrastructure planning
- Informs climate-sensitive urban design and mitigates urban heat islands
- Enables efficient irrigation and resource use in agriculture through fine-grained predictions

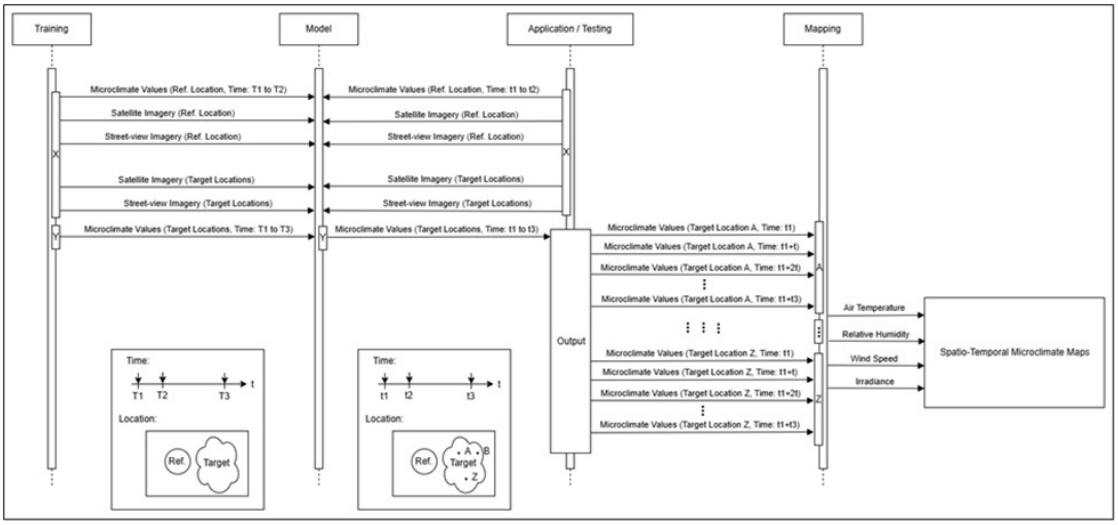
## 2.3 Future Work

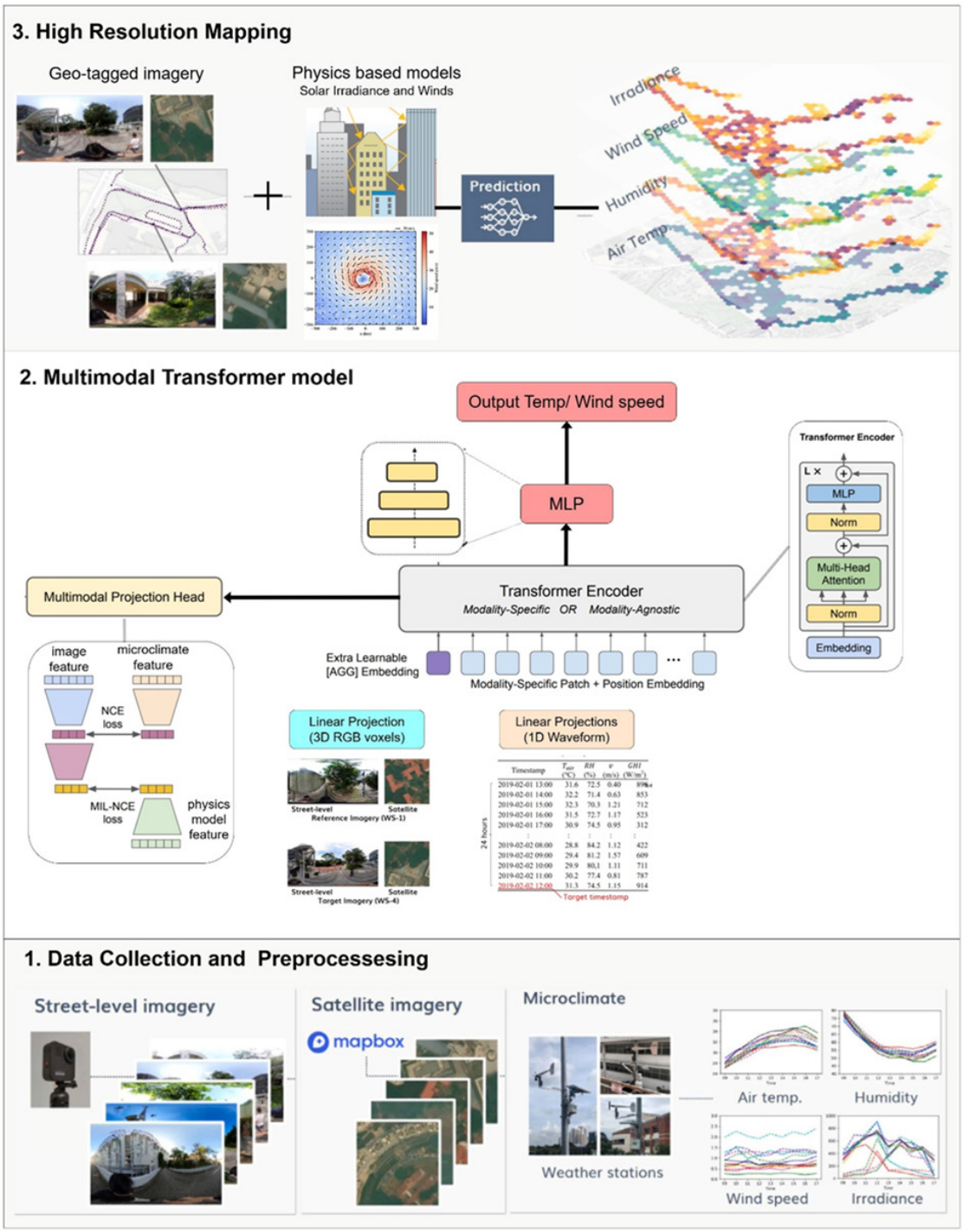
- Data Expansion: Increase training data with seasonal and multi-location inputs
- Model Enhancement: Integrate multi-modal transformer models for better prediction
- Real-Time Optimization: Streamline model for deployment on edge or real-time systems
- Collaborations: Explore partnerships with urban authorities and agritech firms for validation and scale-up
- Infrastructure Needs: Additional sensors, high-res imagery access, and GPU upgrades

## 3 Use Case Requirements

- **REQ-01**: Access to satellite imagery APIs (e.g., Mapbox Static Tiles) for geo-tagged visuals
- **REQ-02**: Distributed weather stations to collect real-time climate parameters
- **REQ-03**: Camera-enabled image collection system near sensor locations
- **REQ-04**: GPU-enabled computing infrastructure for model training and deployment

## 4 Sequence Diagram






## 5 References


[1] Han, J., et al. (2024). Microclimate spatio-temporal prediction using deep learning and land use data. Building and Environment.

[2] Fujiwara, K., et al. (2024). Microclimate Vision: Multimodal prediction of climatic parameters using street-level and satellite imagery. Sustainable Cities and Society.

[3] Zanchi, M., et al. (2023). Harnessing deep learning to forecast local microclimate using global climate data. Sci Rep, 13, 21062.

[4] Deznabi, I., et al. (2024). Zero-shot Microclimate Prediction with Deep Learning. arXiv:2401.02665

[5] Mapbox Satellite Imagery

[6] Google Street View

[7] https://www.sciencedirect.com/science/article/pii/S2210670724005584

[8] https://github.com/syed-azim-git/Microcimate-Prediction

# Use Case 20: AI-Based Real-Time Elephant Monitoring and Human Elephant Conflict Mitigation System

**Organization**: Mbeya University of Science and Technology (MUST)

**Country**: Tanzania

**Contact Person(s)**:

**Primary Contact**: Majaliwa Reuben Joseph – majaliwajoseph33@gmail.com

Secondary Contact: +255752293163

Additional Contact: +255787645162

## 1 Use Case Summary Table

| Item | Details |
|---|---|
| Category | Environmental Protection, Wildlife Monitoring, Conflict Mitigation |
| Problem Addressed | Human-elephant conflicts (HEC) in rural Tanzania lead to crop damage, property loss, and human injuries. Traditional methods are reactive and ineffective. |
| Key Aspects of Solution | – GPS tracking of elephants<br>– AI-based geofencing<br>– GSM alert system<br>– Automated deterrents<br>– Behavioral prediction through historical data |
| Technology Keywords | AI, Geofencing, GPS, GSM, IoT, Wildlife Conservation |
| Data Availability | Private (GPS tracking data, danger zone maps, behavior logs) |
| Metadata (Type of Data) | GPS coordinates, timestamps, zone entries, movement patterns |
| Model Training and Fine-Tuning | Rule-based geofencing; future integration of ML for predictive behavioral analysis |
| Testbeds or Pilot Deployments | Planned deployments in elephant corridors near Tanzanian farms and villages |

# 2 Use Case Description

## 2.1 Description

Context

As elephant habitats shrink due to agricultural expansion, elephants increasingly interact with human settlements, leading to dangerous and costly conflict. These events damage crops and properties, cause injuries, and endanger elephant populations due to retaliatory actions.

Objectives

This use case aims to reduce human-elephant conflict through a smart, proactive AI-driven solution that:

1. Tracks elephants using GPS
2. Monitors proximity to danger zones via AI-based geofencing
3. Sends real-time alerts to communities via GSM
4. Activates non-lethal deterrents
5. Logs data for long-term behavioral analysis and system improvement

Innovation

The solution integrates GPS collars, GSM alerts, AI logic, and smart deterrents such as buzzers and strobes. A microcontroller processes data in real time, and future upgrades will introduce machine learning models for predicting elephant behavior based on logged data.

Expected Impact

- Protects crops and properties through early intervention
- Enhances community safety
- Reduces retaliatory killings, aiding conservation
- Enables data-driven strategies for wildlife authorities
- Scales to other wildlife conflict scenarios

**Use Case Status**: Prototype developed; deployment planned in wildlife corridors

### Partners

1. Mbeya University of Science and Technology (MUST) – Lead on hardware and AI system development
2. Tanzania Wildlife Research Institute (TAWIRI) – Wildlife access, GPS collaring, ethical compliance
3. Tanzania National Parks Authority (TANAPA) – Operational support within national parks
4. Local Government Authorities – Community mobilization, deployment facilitation, awareness creation

## 2.2 Benefits of use case

- Conserves biodiversity and prevents habitat degradation through reduced conflict

- Enhances resilience to wildlife-related climate stressors
- Leverages AI and IoT to build sustainable, community-centered monitoring systems

## 2.3 Future Work

**Next Phases**:

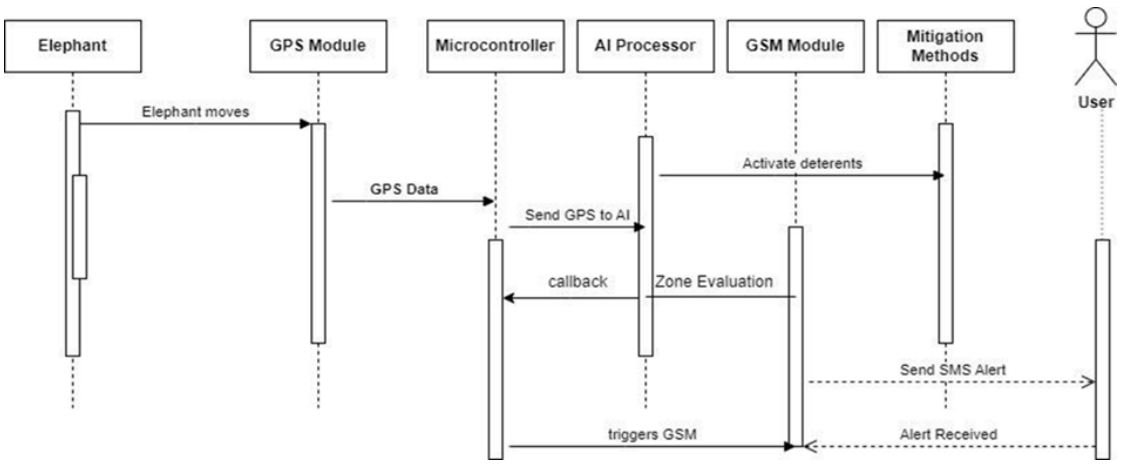
- Develop machine learning models for predicting future conflicts
- Add smart deterrents (e.g., water cannons, drone surveillance)
- Build mobile app for real-time visualization and interaction
- Deploy admin dashboard for centralized monitoring
- Expand GPS collar deployments and cloud infrastructure
- Integrate new datasets (crop seasons, weather, land use)
  Launch wider training programs for communities and rangers

Adapt the platform to other species (e.g., buffalo, hippos)

# 3 Use Case Requirements

- **REQ-01**: GPS collars for real-time elephant tracking
- **REQ-02**: AI-powered geofencing to identify high-risk zone entry
- **REQ-03**: GSM modules to send alerts and control deterrent systems
- **REQ-04**: Microcontrollers (Arduino UNO, ESP32) for local processing
- **REQ-05**: Deterrent devices (flashing lights, bee sounds)
- **REQ-06**: Solar or battery-powered units for off-grid operation
- **REQ-07**: SD card storage for logging movement data
- **REQ-08**: Mobile interface for alert access by local users
- **REQ-09**: Community training and awareness programs
- **REQ-10**: Future ML module for predictive conflict prevention

# 4 Sequence Diagram



## 5 References

[1] Groom, R.J. & Western, D. (2013). *Oryx, 47(1), 34-40*.

[2] Fernando, P. et al. (2005). Biodiversity and Conservation, 14(10), 2465–2481.

[3] World Bank (2022). Community-Based Wildlife Conservation Toolkit.

[4] TAWIRI Reports (2021–2023).

[5] United Nations SDGs – https://sdgs.un.org/goals

[6] Arduino Official Documentation – https://www.arduino.cc/en/Guide

[7] ITU-T Y.2060 (06/2012) – IoT Overview

[8] Miller, S.M. et al. (2016). Conservation Biology, 30(3), 593–598

# Use Case 21: Driven Meteorological Modeling: Transforming Wind Speed and Direction Prediction

**Organization**: Environmental Research Institute, University of Sadat City & SAMI Advanced Electronics

**Division**: Meteorology and Environment

**Country**: Saudi Arabia

**Contact Person(s)**:

**Name**: Mohamed Azouz

**Email**: azouzster@gmail.com

**Phone**: +966504354224

## 1 Use Case Summary Table

| Item | Details |
|---|---|
| Category | Energy Efficiency, Climate Action, Sustainability |
| Problem Addressed | Traditional NWP models are computationally intensive and struggle to deliver accurate, real-time wind forecasts – critical for renewable energy and climate resilience. |
| Key Aspects of Solution | – AI-enhanced forecasting using LSTM, CNN, RNN<br>– Integration with ECMWF's AIFS<br>– Real-time and historical wind data from sensors, satellites, and weather stations<br>– Improved grid stability and renewable energy output |
| Technology Keywords | Artificial Intelligence, Machine Learning, Wind Prediction, Deep Learning, Smart Grid, AIFS, Meteorological Modeling |
| Data Availability | Public and Private:<br>Public sources include ECMWF, NOAA, NASA, Copernicus<br>Private station and IoT sensor data |
| Metadata (Type of Data) | Time-series wind speed and direction data, satellite observations, real-time sensor feeds |
| Model Training and Fine-Tuning | Deep learning models trained on multi-source wind datasets; integrated with AIFS for global modeling robustness |
| Testbeds or Pilot Deployments | Proof-of-concept under development; focus regions include the Red Sea coast (Saudi Arabia) and Gulf of Suez (Egypt) |

(continued)

| Item | Details |
|---|---|
| Repository URLs | ECMWF Data Portal, Copernicus Climate Data Store, NASA GMAO, NOAA, IRENA Data |

# 2 Use Case Description

## 2.1 Description

Accurate wind forecasting plays a pivotal role in optimizing renewable energy deployment and balancing smart grid operations. However, traditional Numerical Weather Prediction (NWP) systems, though effective on a macro scale, face limitations in real-time performance, spatial resolution, and computational cost.

This AI-driven use case addresses those limitations by integrating deep learning models–LSTM, CNN, and RNN–into wind speed and direction forecasting pipelines. These models are trained on real-time and historical datasets from IoT sensors, meteorological stations, and satellite feeds. The project is also aligned with ECMWF's AI-based Forecasting System (AIFS) to strengthen predictions using global atmospheric modeling.

Benefits & Applications:

- Optimized Wind Farm Operations: Accurate forecasts improve turbine scheduling, increasing efficiency and reducing backup energy reliance.
- Smart Grid Integration: Forecasts help stabilize electricity supply-demand balance, essential for high-penetration renewable systems.
- Climate Risk Mitigation: Enables early warning for severe wind events (e.g., dust storms, cyclones).
- Policy & Planning: Informs infrastructure investment, regional energy strategies, and SDG-compliant climate planning.

**Use Case Status**: Part of a larger ongoing research initiative in Saudi Arabia and Egypt.

### Partners

- Environmental Research Institute, University of Sadat City
- SAMI Advanced Electronics – Research and technology integration

## 2.2 Benefits of use case

- Optimized Wind Farm Operations: Accurate forecasts improve turbine scheduling, increasing efficiency and reducing backup energy reliance.
- Smart Grid Integration: Forecasts help stabilize electricity supply-demand balance, essential for high-penetration renewable systems.
- Climate Risk Mitigation: Enables early warning for severe wind events (e.g., dust storms, cyclones).
- Policy & Planning: Informs infrastructure investment, regional energy strategies, and SDG-compliant climate planning.

## 2.3 Future Work

**Proof of Concept**:

- Build a prototype integrating AIFS and custom deep learning models.
- Validate against historical benchmarks and compare with conventional NWP performance.

**Standardization**:

- Define KPIs for forecast performance and system efficiency.
- Collaborate with ITU and WMO to create global AI forecasting guidelines.
- Develop a framework for scalable deployment in smart grids and national weather services.

**Why It Matters**:

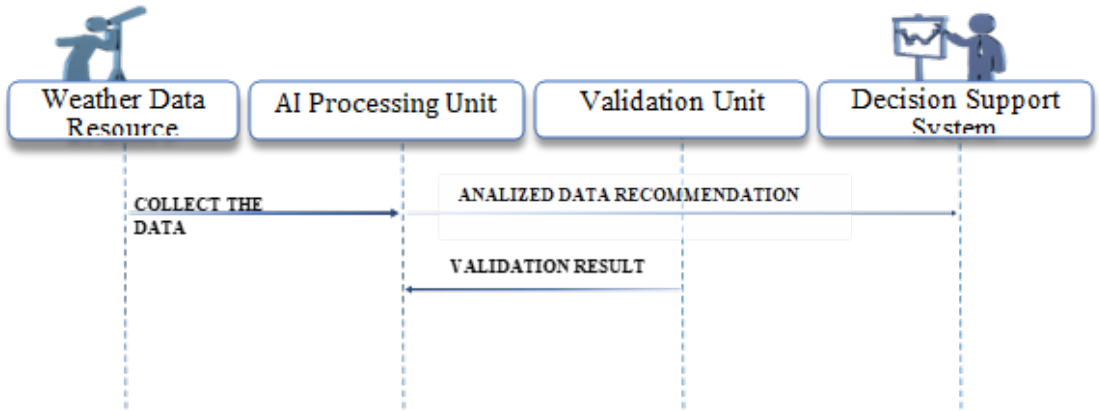
- Energy Sector: Supports clean energy expansion with minimal waste.
- Climate Resilience: Enhances disaster preparedness and reduces wind-related losses.

**Global Leadership**: Builds AI-based meteorological innovation capacity in the MENA region and beyond.

# 3 Use Case Requirements

- **REQ-01**: Access to multi-source meteorological data (satellites, stations, sensors)
- **REQ-02**: Deep learning framework (LSTM, CNN, RNN) integration pipeline
- **REQ-03**: Cloud-based infrastructure for training and real-time inference
- **REQ-04**: Standardized accuracy benchmarks for comparing AI vs NWP results
- **REQ-05**: Partnership with ECMWF and Copernicus to leverage AIFS
- **REQ-06**: Open-access visualization tools for policymakers and energy operators
- **REQ-07**: Disaster alert integrations for extreme wind events
- **REQ-08**: Regulatory alignment with ITU and WMO standards
- **REQ-09**: AI explainability components for transparency and trust
- **REQ-10**: Stakeholder training and outreach to utility providers and planners

# 4 Sequence Diagram



## 5 References

[1] ECMWF AI Forecasts Become Operational (2025)

[2] Royal Meteorological Society – Advances in AI Weather Forecasting

[3] ITU AI for Good Initiative

[4] NOAA National Centers for Environmental Information

[5] GMAO/NASA Global Modelling Data

[6] IRENA Renewable Energy Data Hub

# Use Case 22: Polar Eco-Carbon Intelligence Platform (PECIP): an AI-Driven ESG Project for Climate Action and Global Sustainability

**Organization:** The Cooperative Open Intelligent Computing Industry Alliance (Joint submission with Polar Research Institute of Hong Kong & Intel)

**Country:** Hong Kong SAR, China

**Contact Person(S)**:

Primary: Jun Wang, service@coia.ai

Secondary: Wei Meng, mengwei@coia.ai

## 1 Use Case Summary Table

| Item | Details |
|---|---|
| Category | Climate Change/Natural Disaster |
| Problem Addressed | Traditional methods of monitoring the polar regions face significant challenges due to the unique characteristics of these environments, as well as the complexity of the processes involved, such as limited accessibility, sparse and inadequate data, high costs, delayed data processing, inadequate early warning systems.<br>We deploy or utilize various data collection sensors in the polar regions to transmit data to an AI platform for predicting climate change in the polar regions. For AI platforms, we use the open source project OPEA (Open Platform for Enterprise AI), deploy PyTorch, and perform enhanced pre training and fine-tuning on open source LLMs (LLaMa or DeepSeek). [4] [5] For real-time data, we use RAG for real-time data synchronization. |

(continued)

| Item | Details |
|---|---|
| Key Aspects of Solution | 1. Real-Time Polar Ecosystem Monitoring<br>AI-driven models provide precise insights into glacier melt, sea ice dynamics, permafrost thaw, and marine ecosystems. These tools enhance early warnings, enabling policymakers to respond to rapid polar changes and mitigate global impacts.<br>2. AI-Enhanced Microalgae Cultivation<br>AI boosts carbon sequestration efficiency by 30%, making microalgae cultivation a scalable and cost-effective solution. This innovation reduces emissions, generates renewable energy, and supports sustainable bioproducts.<br>3. Open AI Platform for Collaboration<br>An open AI platform democratizes data access, fosters interdisciplinary research, and accelerates global climate action. By enabling real-time collaboration, this platform empowers decision-makers with actionable insights.<br>4. ESG and Education<br>In partnership with The Education University of Hong Kong, PECIP develops "AI+ESG" courses and policy toolkits, driving corporate green transformation.<br>5. Cross-Domain Expansion<br>Polar technologies are applied to urban clusters in the Guangdong-Hong Kong-Macau Greater Bay Area, supporting smart water management and low-carbon development. |
| Technology Keywords | Large Language Model, Retrieval-augmented Generation, Deep Neural Networks |
| Data Availability | Public : Polar satellite remote sensing data, partially open. E.g. NASA Earthdata, European Space Agency (ESA) Copernicus Open Access Hub, NOAA CLASS (Comprehensive Large Array-Data Stewardship System). [6] [7] [8]<br>Private: Arctic "Purple Jade Station", pole station of PRIHK, expedition data. |
| Metadata (Type of Data) | Text and visual |

(continued)

| Item | Details |
|---|---|
| Model Training and Fine-Tuning | For AI Training, we use the open source project OPEA (Open Platform for Enterprise AI), deploy PyTorch, and perform enhanced pre training and fine-tuning on open source LLMs (LLaMa or DeepSeek). For real-time data, we use RAG for real-time data synchronization.<br>We use the following technologies to improve the training efficiency:<br>3D parallel acceleration: it can reduce the requirements for video memory and computing power on a single device, and combined with distributed technology, enables large model training on a large number of devices.<br>ZeRO video memory optimization: it performs partition optimization on static and dynamic storage during training, reducing the model's demand for video memory, and the calculation and transmission time can be over-lapped.<br>FlashAttention: reduce data transfer between Static Random-Access Memory (SRAM) and High Bandwidth Memory (HBM) improves training speed through calculation blocks and operator fusion.<br>We use Low-Rank Adaptation (LoRA) for fine tuning. It adds additional low-rank matrices and train only these low-rank matrices. |
| Testbeds or Pilot Deployments | https://innoport.cuhk.edu.hk/single-cubiczine/hokinchung/ [9]<br>https://github.com/opea-project [10]<br>The AI platform and tools were provided by Intel above. |

# 2 Use Case Description

## 2.1 Description

The initiator of the project, Prof. Ho Kin-Chung BBS, is a world-renowned environmental protection expert. He has conducted scientific expeditions to the Arctic and Antarctic several times. COIA joins forces with Professor Ho to build AI-driven ESG Project for Climate Action and Global Sustainability.

Ecological changes in the polar regions act as a catalyst for global climate change, amplifying warming and disrupting Earth's delicate balance. Immediate action to mitigate greenhouse gas emissions and protect polar ecosystems is essential to limit these cascading effects and safeguard Earth's climate for future generations.

Problems and Limitations:

Traditional methods of monitoring the polar regions face significant challenges due to the unique characteristics of these environments, as well as the complexity of the processes involved.

1) Limited Accessibility

The harsh weather, extreme cold, and remoteness of the polar regions make it difficult to deploy and maintain monitoring equipment such as weather stations, buoys, and other sensors.

Physical access to these regions, especially during winter months, is often impossible, leading to gaps in data collection.

2) Sparse and Inadequate Data

Monitoring networks in polar regions are sparse compared to those in more accessible areas. This results in significant gaps in data coverage for both the Arctic and Antarctic.

Satellite data, while useful, often lacks the resolution or precision needed to monitor smaller-scale ecological or climatic changes.

3) High Costs

Deploying and maintaining monitoring equipment in the polar regions is expensive due to the logistical challenges and the need for specialized equipment designed to withstand extreme conditions.

Research expeditions require significant financial and human resources, limiting the frequency and geographic reach of such efforts.

4) Delayed Data Processing

Traditional data collection methods often rely on manual or semi-automated processing, which can result in delays in interpreting and sharing critical findings.

This limits the ability to provide timely early warnings for events like ice sheet collapse, permafrost thaw, or extreme weather.

5) Inadequate Early Warning Systems

Existing early warning systems may not be capable of forecasting extreme events in polar regions (e.g., rapid ice shelf disintegration) with sufficient accuracy or lead time.

Systems often lack the ability to integrate diverse datasets from different sources (e.g., satellite imagery, ocean sensors, and climate models).

PECIP integrates polar ecological monitoring, microalgae carbon capture, and ESG (Environmental, Social, Governance) policy research to drive climate governance, carbon neutrality, and cross-border collaboration through AI, advancing the synergistic achievement of the UN SDGs. It uses real-time data and offline data. This hybrid approach ensures comprehensive data coverage, combining the precision of historical data with the immediacy of real-time inputs. It implements the following functions:

1) Real-time analysis of polar ecosystems via AI models (e.g., glacier melt prediction)

   AI models enable real-time analysis of polar ecosystems, providing unprecedented insights into glacier melt, sea ice dynamics, permafrost thaw, and marine ecosystems. These technologies enhance predictive accuracy, detect subtle changes, and enable early warnings for critical environmental events. By integrating AI into polar monitoring, scientists and policymakers can better understand and respond to the rapid changes occurring in these fragile ecosystems and their profound impacts on Earth's climate system.

2) AI-driven microalgae cultivation increases carbon absorption by 30%

   AI-driven microalgae cultivation represents a breakthrough in carbon sequestration, increasing $CO_2$ absorption efficiency by 30% while addressing scalability, cost, and resource challenges. By integrating advanced AI technologies, this approach offers a sustainable and multi-functional solution to combat climate change. With further research and investment, it holds great promise for reducing industrial emissions, generating

renewable energy, and creating valuable bioproducts, contributing to a more sustainable future.

3) Open AI platform enables global data sharing and collaborative research

   An open AI platform for global data sharing and collaborative research represents a game-changing approach to addressing the challenges of polar monitoring and climate change. By enabling real-time collaboration, democratizing access to advanced tools, and fostering interdisciplinary research, such platforms accelerate scientific discovery and empower decision-makers with actionable insights. As climate challenges become more urgent, open AI platforms will play an increasingly vital role in uniting the global community to protect Earth's most fragile ecosystems.

4) **ESG Institute Collaboration:** Leveraging resources from The Education University of Hong Kong to develop "AI+ESG" educational courses and policy toolkits, empowering corporate green transformation.

5) **Cross-Domain Expansion:** Applying polar technologies to the Guangdong-Hong Kong-Macau Greater Bay Area urban clusters, supporting smart water management and low-carbon community development.

## 2.2 Benefits of use case

The integration of AI and LLMs (e.g., machine learning models, climate-specific AI models, and large language models) offers transformative potential for polar monitoring and early warning systems. These technologies address many of the limitations of traditional methods by enhancing data collection, analysis, and prediction capabilities. Key benefits include:

1) Remote Sensing and Autonomous Systems

   AI can enhance the capabilities of remote sensing technologies (e.g., satellites, drones, and autonomous underwater vehicles) by automating data collection, image analysis, and anomaly detection.

   Autonomous systems, equipped with AI, can operate in harsh polar conditions without human intervention, collecting high-resolution data at lower costs.

2) Improved Data Integration and Analysis

   AI can process and integrate vast amounts of heterogeneous data, including satellite imagery, sensor data, historical records, and real-time observations, which would be impossible for humans to analyze manually.

   Machine learning models can identify patterns, trends, and anomalies that may not be apparent through traditional statistical methods.

3) Cost Efficiency and Scalability

   AI-driven monitoring systems reduce reliance on expensive field expeditions by maximizing the utility of remote sensing and automated data collection.

   Once developed, AI models can be scaled to monitor large areas of the polar regions without a proportional increase in costs.

4) Real-Time Monitoring

   AI systems can analyze data in real-time, enabling faster detection of critical changes, such as sudden ice shelf collapses or new cracks in glaciers.

5) Early Warning

   Early warning systems powered by AI can provide actionable insights to governments, researchers, and communities, allowing for timely responses to potential disasters.

6) Scenario Simulation and Decision Support

Large models can simulate a wide range of scenarios, helping policymakers and researchers understand the potential impacts of specific events (e.g., ice sheet collapse or methane release) and make informed decisions.

AI-powered decision support tools can recommend optimal intervention strategies to mitigate risks associated with polar changes.

The PECIP platform creates a synergistic network across sustainable development by integrating technological innovation and global collaboration. Ecological changes in the polar regions act as a catalyst for global climate change, amplifying warming and disrupting Earth's delicate balance. Immediate action to mitigate greenhouse gas emissions and protect polar ecosystems is essential to limit these cascading effects and safeguard Earth's climate for future generations. Climate Action is directly advanced through AI-powered polar monitoring systems, which track glacier dynamics and marine pollution in real-time, providing critical data for climate models and marine conservation policies.

The platform's broader societal impact lies in bridging Quality Education. By partnering with the ESG Institute, PECIP delivers AI-driven training programs and open-access tools, equipping youth with green skills for sustainable employment in sectors like carbon finance and eco-engineering. Modular technologies (e.g., microalgae systems) enable infrastructure upgrades in developing nations, while federated learning frameworks foster South-South data sharing. For instance, Brazilian researchers utilize PECIP's open polar datasets to optimize Amazon rainforest conservation, demonstrating how localized solutions emerge from global cooperation.

Crucially, all components are anchored in ethical governance – AI transparency ensures accountability in carbon credit markets, while participatory platforms allow citizens to engage in climate action. This multidimensional approach not only aligns with UN priorities but redefines sustainability as a collective endeavor where technology, education, and equity converge.

**Use Case Status:** Research Phase, implementing in lab of The Education University of Hong Kong

**Partners:** Polar Research Institute of Hong Kong; Intel

The integration of AI and LLMs (e.g., machine learning models, climate-specific AI models, and large language models) offers transformative potential for polar monitoring and early warning systems. These technologies address many of the limitations of traditional methods by enhancing data collection, analysis, and prediction capabilities. Key benefits include:

**1. Remote Sensing and Autonomous Systems**

AI can enhance the capabilities of remote sensing technologies (e.g., satellites, drones, and autonomous underwater vehicles) by automating data collection, image analysis, and anomaly detection.

Autonomous systems, equipped with AI, can operate in harsh polar conditions without human intervention, collecting high-resolution data at lower costs.

**2. Improved Data Integration and Analysis**

AI can process and integrate vast amounts of heterogeneous data, including satellite imagery, sensor data, historical records, and real-time observations, which would be impossible for humans to analyze manually.

Machine learning models can identify patterns, trends, and anomalies that may not be apparent through traditional statistical methods.

**3. Cost Efficiency and Scalability**

AI-driven monitoring systems reduce reliance on expensive field expeditions by maximizing the utility of remote sensing and automated data collection.

Once developed, AI models can be scaled to monitor large areas of the polar regions without a proportional increase in costs.

**4. Real-Time Monitoring**

AI systems can analyze data in real-time, enabling faster detection of critical changes, such as sudden ice shelf collapses or new cracks in glaciers.

**5. Early Warning**

Early warning systems powered by AI can provide actionable insights to governments, researchers, and communities, allowing for timely responses to potential disasters.

**6. Scenario Simulation and Decision Support**

Large models can simulate a wide range of scenarios, helping policymakers and researchers understand the potential impacts of specific events (e.g., ice sheet collapse or methane release) and make informed decisions.

AI-powered decision support tools can recommend optimal intervention strategies to mitigate risks associated with polar changes.

The PECIP platform creates a synergistic network across sustainable development by integrating technological innovation and global collaboration. Ecological changes in the polar regions act as a catalyst for global climate change, amplifying warming and disrupting Earth's delicate balance. Immediate action to mitigate greenhouse gas emissions and protect polar ecosystems is essential to limit these cascading effects and safeguard Earth's climate for future generations. Climate Action is directly advanced through AI-powered polar monitoring systems, which track glacier dynamics and marine pollution in real-time, providing critical data for climate models and marine conservation policies.

The platform’s broader societal impact lies in bridging Quality Education. By partnering with the ESG Institute, PECIP delivers AI-driven training programs and open-access tools, equipping youth with green skills for sustainable employment in sectors like carbon finance and eco-engineering. Modular technologies (e.g., microalgae systems) enable infrastructure upgrades in developing nations, while federated learning frameworks foster South-South data sharing. For instance, Brazilian researchers utilize PECIP’s open polar datasets to optimize Amazon rainforest conservation, demonstrating how localized solutions emerge from global cooperation.

Crucially, all components are anchored in ethical governance—AI transparency ensures accountability in carbon credit markets, while participatory platforms allow citizens to engage in climate action. This multidimensional approach not only aligns with UN priorities but redefines sustainability as a collective endeavor where technology, education, and equity converge.

## 2.3 Future Work

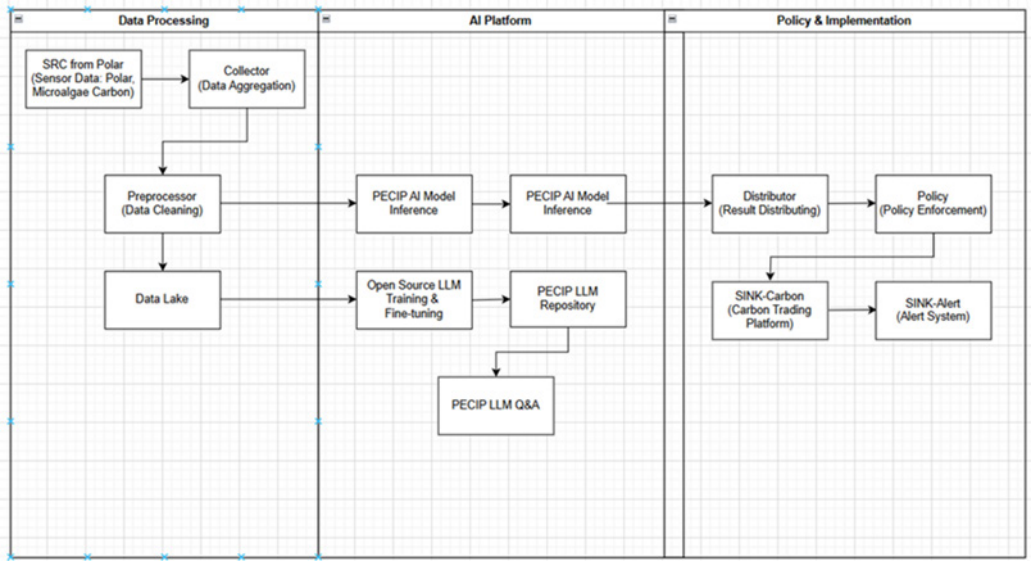
- **Improved Predictive Accuracy:** Continuous refinement of AI models for polar ecosystem analysis, such as better glacier melt predictions, enhanced sea ice monitoring, and more precise permafrost thaw simulations.
- **Integration of Multimodal Data**: Develop AI models capable of integrating a broader range of data types, such as satellite images, drone footage, IoT sensor data, and historical climate records, to improve the granularity and accuracy of insights.
- **Real-time AI Optimization:** Enhance real-time capabilities to detect subtle changes and respond faster to critical environmental events; Development of Advanced Microalgae Cultivation Models
- **AI-Driven Optimization**: Further develop AI-based cultivation systems to improve scalability, maximize carbon absorption, and minimize resource consumption.
- **Integration with Renewable Energy Systems:** Implement renewable energy-powered microalgae cultivation systems to ensure the process is sustainable and cost-effective.
- **Global ESG Standards Alignment**: Align the platform with globally recognized ESG frameworks, such as the Task Force on Climate-Related Financial Disclosures (TCFD) and ITU standards for environmental management.
- **Open Data Standards:** Establish standardized protocols for aggregating, processing, and sharing data across nations and organizations to foster global collaboration.
- **Open Data Standards:** Launch a challenge based on the OPEA platform during the July workshop or following events.

# 3 Use Case Requirements

- **REQ-01:** It is required to have advanced AI models for real-time polar ecosystem analysis and predictions (glacier melt, sea ice, permafrost).
- **REQ-02**: It is required to have AI-optimized microalgae cultivation systems for enhanced carbon capture (30% efficiency).
- **REQ-03**: It is required to have autonomous monitoring systems (drones, underwater vehicles) resilient to polar conditions.
- **REQ-04**: It is required to have high-resolution remote sensing integration with AI analysis (satellites, drones).
- **REQ-05**: It is required to have real-time data processing infrastructure (edge/cloud computing).
- **REQ-06**: It is required to have scalable cloud storage and secure APIs for global data sharing.
- **REQ-07**: It is required to have predictive early warning systems for environmental anomalies (ice collapse, methane leaks).
- **REQ-08**: It is recommended to have simulation tools for climate scenarios and policy impact assessment.
- **REQ-09:** It is recommended to have Open AI platform with collaborative tools (real-time access, federated learning).

- **REQ-10**: It is recommended to have ESG education and policy development suite (courses, toolkits, corporate support).
- **REQ-11**: It is recommended to have cross-domain application interfaces for urban adaptation (smart water, low-carbon communities).
- **REQ-12**: It is recommended to have international collaboration frameworks (data sharing, governance, standards).
- **REQ-13**: It is required to have deployment and maintenance protocols for polar equipment (logistics, rugged hardware).
- **REQ-14**: It is required to have resource optimization systems (cost, energy, sustainable logistics).
- **REQ-15**: It is required to have data validation and model accuracy verification processes.
- **REQ-16**: It is recommended to have cybersecurity measures and ethical AI guidelines (encryption, bias audits).
- **REQ-17**: It is required to have renewable energy-powered infrastructure (data centers, sensors).
- **REQ-18**: It is recommended to have integration with industrial emission sources for microalgae deployment.
- **REQ-19**: It is recommended to have interactive dashboards and visualization tools for policymakers.
- **REQ-20**: It is required to comply with global environmental and data regulations.

## 4 Sequence Diagram




1) Data Processing

This stage focuses on collecting, cleaning, and storing data for further analysis and AI model training.

### SRC-Polar (Sensor Data: Polar, Microalgae Carbon)

Sensor data from polar ecosystems (e.g., glacier dynamics, sea ice, permafrost) and microalgae carbon monitoring is generated.

### Collector (Data Aggregation)

Data from SRC-Polar is aggregated into a centralized system for further processing.

### Preprocessor (Data Cleaning)

The aggregated data is cleaned and preprocessed to remove errors, inconsistencies, or irrelevant information, ensuring quality input for downstream tasks.

### Data Lake

The cleaned and processed data is stored in a data lake, a centralized repository that supports both structured and unstructured data. It serves as a source for AI model training, fine-tuning, and inference.

2) AI Platform

This stage involves AI model development, training, inference, and the use of large language models (LLMs).

### PECIP AI Model Inference

AI models use the processed data to make inferences about polar ecological changes, microalgae carbon sequestration, and other key metrics.

These inferences provide real-time insights or predictions (e.g., glacier melt rates, carbon absorption efficiency).

### Open Source LLM Training & Fine-Tuning

Large language models (LLMs) are trained and fine-tuned using data from the data lake.

This process ensures the LLMs are optimized for specific tasks, such as answering questions about the polar environment or recommending sustainability practices.

### PECIP LLM Repository

The LLMs are stored in a repository for reuse in various applications, such as Q&A systems or generating reports.

### PECIP LLM Q&A

A specialized system enables users to ask questions and receive answers powered by the fine-tuned LLMs.

3) Policy & Implementation

This stage applies the insights and predictions generated in the earlier stages to real-world climate action and sustainability policies.

### SINK-Carbon (Carbon Trading Platform)

AI-driven insights about carbon sequestration are channeled into a carbon trading platform to support carbon neutrality goals.

This platform enables organizations to trade carbon credits based on measurable reductions in carbon emissions or increases in carbon absorption.

### SINK-Alert (Alert System)

An alert system provides timely warnings about critical polar environmental events (e.g., ice shelf collapses).

These alerts enable rapid responses from researchers, policymakers, and stakeholders.

### Distributor (Result Distributing)

Processed insights and alerts are distributed to relevant stakeholders, such as governments, corporations, and research institutions.

### Policy (Policy Enforcement)

Insights and recommendations are used to enforce or guide policies related to climate governance, carbon neutrality, and sustainability.

## 5 References


[1] W. Fowler, J. Segovia, and C. Ah-Siong, "Moving from OpenAI to Open Source with OPEA White Paper." Accessed: Jun. 09, 2025. [Online]. Available: https://cdrdv2-public.intel.com/847353/moving-from-openai-to-open-source-with-opea-whitepaper.pdf

[2] K. C. Ho, "Overview of harmful algal blooms (red tides) in Hong Kong during 1975–2021," Journal of Oceanology and Limnology, vol. 40, no. 6, pp. 2094–2106, Nov. 2022, doi: https://doi.org/10.1007/s00343-022-2205-z .

[3] H.-C. Kim, "Four Decades of Polar Research in Remote Sensing: A Comprehensive Review," Korean Journal of Remote Sensing, vol. 40, no. 5–2, pp. 867–879, Oct. 2024, doi: https://doi.org/10.7780/kjrs.2024.40.5.2.12.

[4] Meta, "Llama 3.2," Meta Llama, 2022. https://www.llama.com/

[5] deepseek, "DeepSeek," Deepseek.com, 2025. https://www.deepseek.com/en

[6] "Earthdata," Nasa.gov, May 10, 2019. https://earthdata.nasa.gov

[7] "Open Access Hub," scihub.copernicus.eu. https://scihub.copernicus.eu

[8] "NOAA's Comprehensive Large Array-data Stewardship System," Noaa.gov, 2025. https://www.class.noaa.gov (accessed Jun. 09, 2025).

[9] "Cubic Zine Aug 2023: Hong Kong people's polar expedition: cross-generational entrepreneurship to research green gold, microalgae expert Ho Kin-chung achieves a win-win situation for environmental protection and business- InnoPort," InnoPort, Jan. 16, 2025. https://innoport.cuhk.edu.hk/single-cubiczine/hokinchung/ (accessed Jun. 09, 2025).

[10] "OPEA [Open Platform for Enterprise AI]," GitHub, Apr. 03, 2025. https://github.com/opea-project

# Use Case 23: AI-Powered Chatbot for Disaster Response

**Organization:** Ethiopian Red Cross, Ethiopian Meteorological Institute

**Country**: Ethiopia

**Contact Person(s)**:

Primary: Thomas Birhanu Aklilu

## 1 Use Case Summary Table

| Item | Details |
|---|---|
| Category | Climate Change/Natural Disaster |
| Problem Addressed | In disaster scenarios, communication systems often break down, limiting access to life-saving information and support. |
| Key Aspects of Solution | – AI-powered multilingual chatbot<br>– Real-time geospatial routing<br>– NLP for crisis reporting and mental health support<br>– Integration with emergency response systems |
| Technology Keywords | NLP, Geospatial AI, Emergency Response, Crisis Chatbot, Multilingual AI |
| Data Availability | Yes – combination of private and public data |
| Metadata (Type of Data) | Location data, Shelter info, Psychological support resources, Missing persons reports |
| Model Training and Fine-Tuning | Localized NLP model training, fine-tuning with domain experts, real-time routing optimization |
| Testbeds or Pilot Deployments | Planned pilot with Ethiopian Red Cross and NGOs |

## 2 Use Case Description

### 2.1 Description

Disasters such as floods and earthquakes can lead to widespread chaos and collapse of traditional communication infrastructures. To address this, the AI-powered disaster response chatbot provides real-time, multilingual, and location-based assistance to individuals and emergency personnel.

By integrating Natural Language Processing (NLP) with geospatial data and real-time datasets, the chatbot enables:

- Guided evacuation based on current conditions
- Locating nearby shelters and essentials
- Crisis communication and mental health support
- Missing persons reporting and follow-up
- Automated alerts and verified updates

The chatbot is accessible via WhatsApp, SMS, Telegram, and voice, ensuring broad usability. The system also supports responder dashboards through API integration and aims to improve situational awareness using AI.

Use Case Status: Ongoing

**Category:** AI for Disaster Response & Climate Resilience

**Partners:** Ethiopian Red Cross, Ethiopian Meteorological Institute

## 2.2 Benefits of use case

- By guiding people to medical assistance and providing psychological support
- Strengthens emergency response systems through AI-driven automation.
- Ensures better preparedness and quick response to urban disasters.
- Uses AI to predict, detect, and mitigate the impact of climate-related disasters.

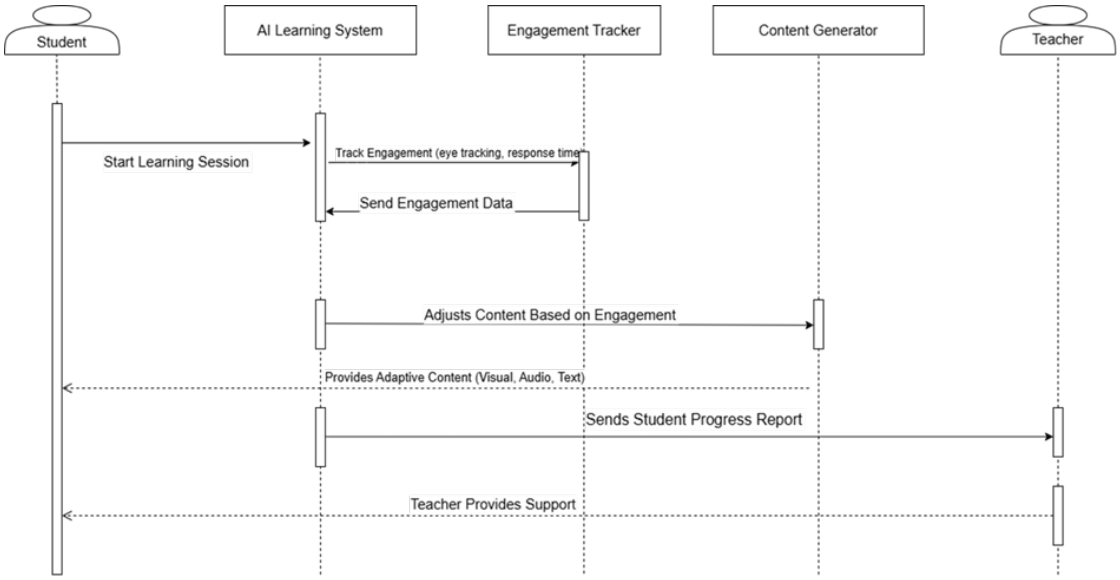
## 2.3 Future Work

- Expansion of local language support and NLP accuracy
- Data collaboration with National Meteorological Agency and Geospatial Institute
- Deployment pilots with humanitarian agencies
- Anonymized dataset publication through GitHub or official repositories
- AI model fine-tuning with disaster response specialists
- Open-source roadmap and interface standardization
- Integration with international alert systems (e.g., GDACS, UN OCHA)

# 3 Use Case Requirements

- **REQ-01**: It is critical to develop NLP models for local languages.
- **REQ-02**: It is expected to integrate with real-time shelter and evacuation data.
- **REQ-03**: It is critical to ensure chatbot access through multiple platforms (SMS, Telegram, WhatsApp).
- **REQ-04**: It is of added value to include voice interface for accessibility.
- **REQ-05**: It is expected to anonymize and geotag data for responder dashboards.
- **REQ-06**: It is of added value to offer psychological support content during emergencies.

## 4 Sequence Diagram




## 5 References

[1] T. Smith et al., "AI-Driven Crisis Communication," *AI for Good Proceedings*, 2023.

[2] IFRC, "Guidelines on Psychological First Aid," 2022.

[3] Ethiopian Red Cross, "Disaster Response Lessons Learned," Internal Report, 2023.

# 4.3. Category 3: 5G

## Use Case 1: AI Full-Process Low-Code Middle Platform Based on the '1+1+3+N' Architecture

**Organization:** China Telecom HongXin Information Technology Co.,Ltd

**Country**: China

**Contact Person(s)**:

Shaoshuai Che, cheshaoshuai.js@chinatelecom.cn

### 1 Use Case Summary Table

| Item | Details |
|---|---|
| Category | 5G |
| Problem Addressed | The widespread application of artificial intelligence remains limited due to high thresholds, including computer power, infrastructure, technical expertise, data quality, and tool complexity, which hinder its integration across industries and sectors. |
| Key Aspects of Solution | A complete Low-Code Middle Platform Based on the '1+1+3+N' Architecture is developed to support the full lifecycle of AI model production using the following components<br>An integrated platform for annotation, training, and inference, covering the full lifecycle workflow of AI model production<br>MaaS model services covering images, text, and videos,<br>Three AI tool platforms include a large model and agent aggregation platform, a visual AI algorithm platform, voice customer service dialog platform.<br>The solution supports both localized deployment & cloud deployment |
| Technology Keywords | All-target recognition, multimodal parsing, multimodal retrieval, multimodal Q&A, big model, cold start, AI agent. |
| Data Availability | Not available |
| Metadata (Type of Data) | Text data, image data, video data |
| Model Training and Fine-Tuning | Model training techniques: small sample robust techniques, GPU acceleration techniques, smart labeling, cold-start and multi-stage training;<br>Base models: DeepSeek, Qwen 2.5, Star Big Model, etc. |

(continued)

| Item | Details |
|---|---|
| Testbeds or Pilot Deployments | The testbed is used internally, and the use case is part of a broader product development effort. |
| Code Repositories | Not available |

# 2 Use Case Description

## 2.1 Description

The development of large models has brought new opportunities for industries, but it has also presented challenges, including a high computing power threshold, increased costs of training and reasoning, difficulty in selecting base models, high technical complexity in deployment, and difficulty in scene adaptation. In order to reduce the technical threshold of AI modeling and application engineering, boost the scale of AI large model application, and promote the process of AI+ action in various industries, this case establishes a set of one-stop modeling and application toolchain platform covering the whole process of AI, providing full-stack services and realizing the efficient sharing of data, arithmetic power, models, tools, and applications.

AI intelligent middle platform, "1+1+3+N" architecture [5]:

- 1 An integrated platform for annotation, training, and inference.
- 2 MaaS model services covering images, text, and videos.
- 3 AI tool platforms (Large Model and Agent Aggregation Platform,Visual AI analysis application platform, Voice customer service application construction platform).
- N applications in 10+ industries, such as transportation, culture and tourism, city, government, and enterprise management.

This use case changes the limitations of existing local code development, such as high technical threshold, low efficiency, poor sharing, etc., and provides AI R&D guidelines and resource sharing through a one-stop platform.

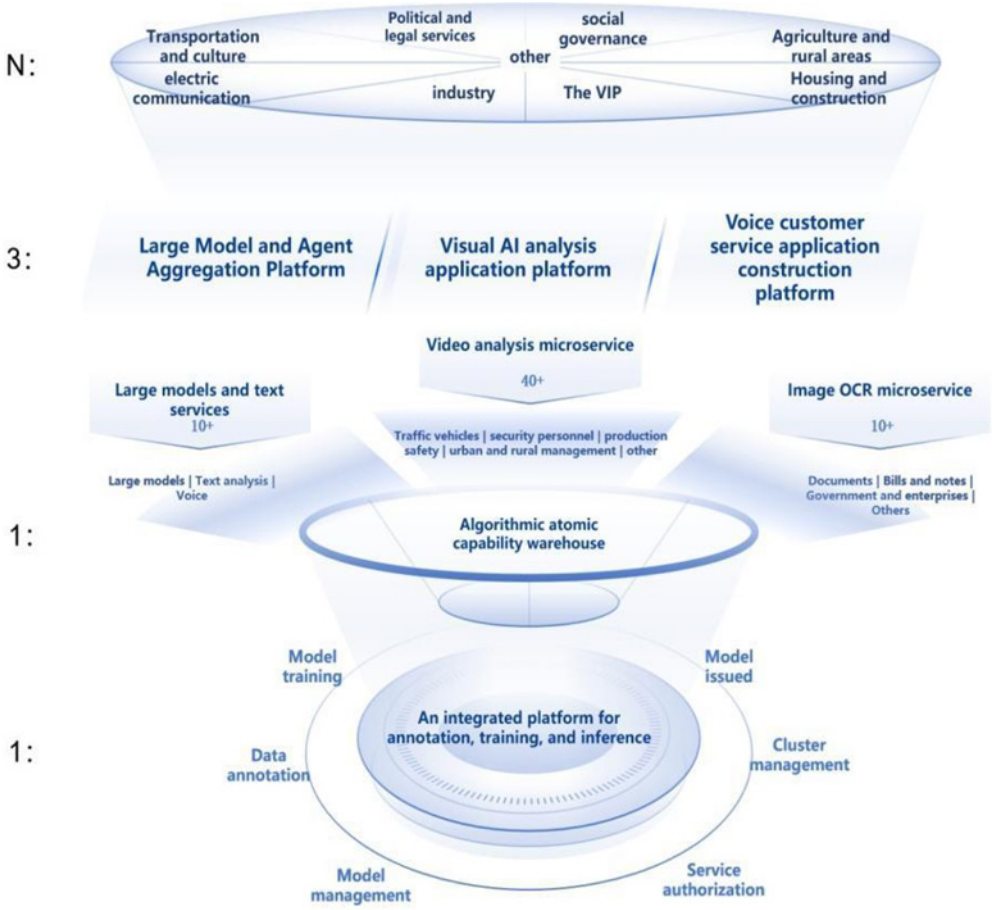



Figure 1: AI Intelligent Middle Platform "1+1+3+N" Architecture Diagram

- 1: An integrated platform for annotation, training, and inference

  The integrated platform for annotation, training, and inference, as a one-stop platform for annotation, training, and reasoning, can achieve five unifications of computing power, data, development, models, and capability usage. It supports the development of algorithms through methods such as managing different computing power, (eg., NVIDIA and Huawei) and visual drag-and-drop interfaces, covering the entire lifecycle workflow of AI model production, thereby lowering the R&D threshold [4].

- 2: MaaS model services covering images, text, and videos

  **Image Services**: There are 16 OCR services, including ID card recognition, social security card image recognition, driver's license recognition, vehicle registration certificate (driving license) recognition, train ticket recognition, etc.

  **Video Services [3]:** There are 47 services, including vehicle detection and attribute recognition, non-motor vehicle detection, vehicle traffic flow analysis, person detection and attribute recognition, face detection and attribute recognition, pedestrian traffic flow analysis, crowd gathering detection, head and shoulder hard hat detection for personnel, etc.

  **Text Services**: There are 17 services, including text dialogue emotion recognition, sensitive word detection in text, short text similarity, entity extraction, and keyword extraction, among others.

- 3: AI tool platforms

  Large Model and Agent Aggregation Platform serves as a connection layer between general-purpose large models and industry applications, aggregating various

open-source large models, managing industry data knowledge bases, plugin tools, workflows, and other components, supporting low-code, visual rapid construction of large model intelligent applications, covering application types such as RAG intelligent dialogue, ChatBI intelligent Q&A, Agent intelligence agents, etc.

Visual AI analysis application platform [2] possesses a rich algorithm library and has completed multi-party computing power adaptation supporting collaborative algorithms for various model sizes. Video Analysis Application: The construction functions include video stream access, task scheduling, alarm push notification, etc.; Algorithm Management and Publication: The construction functions include algorithm management, registration, updates, etc.; Image and Video Algorithm Repository: Provides 47 algorithms for four major scenarios including traffic vehicles, security personnel, workplace safety, and urban-rural management.

Voice customer service application construction platform recognizes citizens' intentions and responds flexibly to citizens' questions, diverting some calls and replacing human agents to provide self-service online Q&A, consultation, and work order acceptance services. Filters sensitive information and refuses inappropriate questions; multi-round dialogue guides users to clarify their intentions; anthropomorphic refinement makes conversations more humane.

- N: N industry application solutions

  The following large models could be used in specific applications and domains such as:

  Transportation: transportation policy and regulations consulting big model, traffic management document generation big model, highway inspection and maintenance big model.

  Tourism: tourism guide model, cultural and museum model, cultural and tourism law enforcement model.

  Digital government: big model of government services, big model of intelligent work order, big model of hotline encyclopaedia, big model of intelligent accompaniment.

  Smart city: housing and construction big model, building big model.

  Social governance: grassroots governance big model, grassroots governance big model-multimodal big model [5].

  Smart agriculture: fishery big model.

Partners

**Nanjing University of Posts and Telecommunications [1]**

## 2.2 Benefits of use case

AI Intelligent Middle Platform, as the infrastructure for AI development, relies on industry application innovation facilities to promote SDGs in transportation, social governance, urban and rural areas, universities and governments, etc. Road disease detection, congestion analysis, and personnel intrusion, to enhance road safety and traffic efficiency; video streaming anomalous event detection, to enhance urban and rural safety; gridders taking pictures of multimodal large model recognition to improve social intelligent governance; river float detection, smoke and fire detection, and garbage dumping detection to reduce resource waste and environmental pollution and optimize energy distribution. The one-stop AI solution promotes the improvement of traffic safety efficiency, safety in urban and rural areas, social governance effectiveness, and environmental protection, facilitates social life in the direction

of greater intelligence and efficiency, and promotes more industries to realize digital and intelligent transformation.

## 2.3 Future Work

Continuously optimize the AI intelligent middle platform, deepen the research of large model technology, improve the development of tools and products, and build an ecology to strengthen the industry, academia, and research.

In terms of products, we have improved the product development process, strengthened the user experience and performance of products, and improved the product promotion model. We have established a standardized process for the implementation of AI agent and formulated a standard specification for the implementation of AI agent scenarios, "three stages and nine processes", which covers the research stage, the AI agent design and implementation stage, and the operation and iteration stage, respectively.

In terms of technology, in the direction of linguistic big models, we have studied big model AI agent technology and established an AI agent platform; in the direction of visual big models, we have established a synergistic process for large and small models, and fully fine-tuned the multimodal big models; in the direction of big model production standardization and training, we have expanded the types of visual big models, multimodal big model applications, and arithmetic support, and built a standardization and training platform integrating big and small models.

# 3 Use Case Requirements

- **REQ-1**: It is critical to establish standard training and push integration, low-code development platform covering the full lifecycle workflow of AI model production.
- **REQ-2**: It is of added value to formulate a standard specification for the implementation of AI agent scenarios, "three stages and nine processes" [4], which covers the research stage, the AI agent design and implementation stage, and the operation and iteration stage, respectively.
- **REQ-3**: It is expected that the platform would reduce the technical threshold of AI modelling and application engineering, boost the scale of AI large model application, and promote the process of AI+ action in various industries, this case establishes a set of one-stop modelling and application toolchain platform covering the whole process of AI, providing full-stack services and realizing the efficient sharing of data, arithmetic power, models, tools, and applications.

## 4 Sequence Diagram

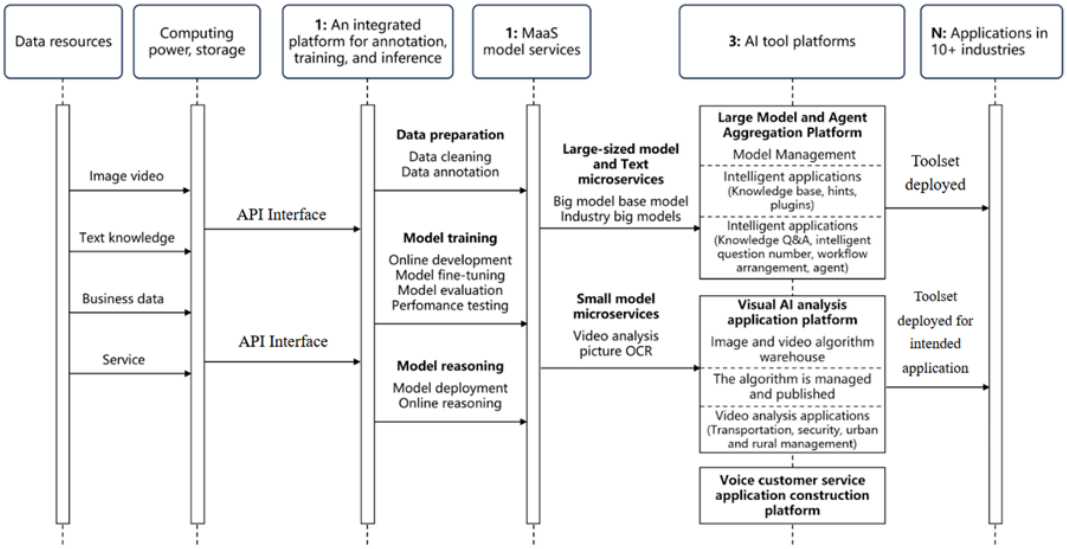



Figure 2 Sequence Diagram of AI Middle Platform Pipeline consisting of the following components:

- **Data phase**: build high-quality datasets, structured data, unstructured data, knowledge graph.
- **Computing power**: integrate smart arithmetic, generalized arithmetic, storage, etc. for training and inference.
- **Standard training and inference stage**: integrate pytorch, LangChian, TensorRT and other frameworks to provide a low-code code development environment for model training and inference.
- **Model service stage**: the model produced by the model, as a service provided to the outside world, including large language models, multimodal models, speech large models, small models, and other components of the MaaS service.
- **Model application construction stage**: Relying on the big model Agent, video toolchain, voice toolchain, and other application toolchains, the model's reasoning ability is used to plan and disassemble the task, and external plug-in tools are used to complete complex tasks[5].
- **Industry application stage**: Industry AI big model applications, such as RAG class applications, ChatBI class applications, and Agent class applications, including transportation big model application assistants, tourism guide application assistants, and grassroots governance assistants.

## 5 References

[1] "Nanjing University of Posts and Telecommunications," [Online]. Available: http://www.njupt.edu.cn/ .

[2] Che Shaoshuai, Zhang Wei, Design of Visual Perception AI Platform Based on Cloud Edge Collaboration [J]. Nanjing: Jiangsu Journal of Communications, April 2024.

[3] Che Shaoshuai, Research on Multi object Detection and Tracking Technology Based on O2O Learning [D]. Nanjing: Nanjing University of Posts and Telecommunications Library, May 2016.

[4] Ministry of Industry and Information Technology (MIIT), Report on the Development of Artificial Intelligence (2024), China Academy of Information and Communications Technology, Beijing, China, 2024, http://www.caict.ac.cn/kxyj/qwfb/bps/202412/t20241210_647283.htm .

[5] Ministry of Industry and Information Technology (MIIT), Blue Book on Artificial Intelligence Governance (2024), China Academy of Information and Communications Technology, Beijing, China, 2024, http://www.caict.ac.cn/kxyj/qwfb/bps/202412/t20241227_649390.htm .

# Use Case 2: Empowering cloud network operation with self-intelligence through the network large model toolchain platform

**Organization**: China Telecommunication Corporation

**Country**: China

**Contact Person(s)**:

Primary contact: Yupei Wang, wangyp6@chinatelecom.cn

Secondary contact: Yungeng Zhang, zhangyg5@chinatelecom.cn

Any other contacts: Feng Wang, wangfeng6@chinatelecom.cn

## 1 Use Case Summary Table

| Item | Details |
|---|---|
| Category | artificial Intelligent |
| Problem Addressed | – Lack of unified AI workflows in telecom operations<br>– Difficulty integrating diverse multimodal telecom data<br>– Limited accessibility of intelligent tools for network teams<br>– Single-model limitations for complex operational needs<br>– High resource costs for model training and deployment<br>– Model training and inference deployment require large amounts of arithmetic resources, which is costly |
| Key Aspects of Solution | – Unified end-to-end toolchain for data, model, and deployment workflows<br>– Built-in multimodal model support for telecom-specific data<br>– Low-code development environment for agent construction<br>– Intelligent orchestration with knowledge bases and plug-ins<br>– Scalable arithmetic resources for AI training and inference |
| Technology Keywords | Large Language Models (LLMs), Network AI Toolchain, Model-as-a-Service (MaaS),Intelligent Agents, Multimodal Data Processing, Cloud Network Management, AI Infrastructure, Automated Network Operations, AI-driven Orchestration, Federated Fine-tuning |
| Data Availability | Private data: we use some sensitive words from https://github.com/jkiss/sensitive-words,<br>but do some processing in the combination with telecom data features. |
| Metadata (Type of Data) | China Telecom's massive network data and knowledge corpora.<br>Text data, Logs, alarms, and configurations |

(continued)

| Item | Details |
|---|---|
| Model Training and Fine-Tuning | Pretrained on open and proprietary corpora, then fine-tuned on multimodal telecom data. Human-in-the-loop supervision ensures task alignment. Uses parameter-efficient tuning, reinforcement learning, and distributed training frameworks. Model performance is tracked via telecom KPIs. |
| Testbeds or Pilot Deployments | we use pilot deployment from http://172.27.221.54:3000/agent-platform-web/login , but do some processing in the combination with telecom data features. |
| Code repositories | Inside China Telecom |

# 2 Use Case Description

## 2.1 Description

As the scale and complexity of network operations continue to grow, traditional AI workflows face challenges such as fragmented development processes, insufficient adaptation to multimodal data, and limited automation in decision-making. Moreover, network-facing intelligent agents often lack standardized, secure, and flexible interfaces to interact with operational systems, limiting their ability to perform real-time diagnostics, control, and service orchestration.

China Telecom has built the industry's entire process toolchain platform that can provide large model dataset processing, training, application construction, and services. The platform can provide multi-modal network professional data processing, powerful computing power support, rich model resources, advanced training and inference methods, and stable training and inference architecture. At the same time, it opens up secure capability interfaces, allowing developers to easily access, customize, and expand large model capabilities to meet diverse needs in different scenarios.

Intelligent agents built through the platform interact with network systems via standardized internal APIs and orchestration frameworks. These interfaces expose functions such as alarm monitoring, ticket issuance, resource querying, configuration adjustment, and service provisioning. By accessing these APIs, agents can autonomously retrieve real-time status, trigger workflows, or execute operational commands under controlled policies, enabling closed-loop automation.

This platform can simplify the entire process of large model corpus processing, training, AI application construction, model deployment and management, greatly improving the widespread application of AI. The platform builds a comprehensive Model-as-a-Service (MaaS) system, which include API, intelligent agent and model deployment, providing customized artificial intelligence solutions more flexibly to meet the needs of different scenarios. Currently, China Telecom are building a 100+ application ecosystem based on toolchains, covering various network fields across the entire network, promoting the transformation and upgrading of cloud network intelligence, and empowering practical scenarios. In the future, we will further focus on researching cross industry customization, technology integration, etc., to provide users with more convenient, efficient, and personalized intelligent experiences, while also promoting the rapid development and innovative upgrading of the entire AI industry.

## 2.2 Benefits of use case

The large model toolchain platform developed by China Telecom provides full-process intelligent services that significantly enhance the quality, efficiency, and sustainability of communication infrastructure.

By targeting the needs of the telecommunications industry, the platform drives innovation through the integration of AI and large model technologies. It collects and utilizes real-time infrastructure data to build intelligent agents that operate across knowledge management, model training, inference, and deployment modules. These agents are tailored to meet specific business requirements, enabling dynamic, data-driven decision-making and operational optimization.

The platform offers advanced tools to process complex telecom data, train domain-specific models, and deploy intelligent agents. It supports a wide range of agent construction methods and allows flexible permission controls for collaborative use across teams or departments. This contributes to more intelligent and efficient cloud network operations, improving resource utilization, reducing operational costs, and addressing challenges in complex telecom scenarios.

By strengthening AI-driven infrastructure management, promoting automation, and fostering industry-level innovation, the platform lays a strong foundation for long-term technological advancement and more resilient, adaptive telecom systems.

## 2.3 Future Work

In 2025, we will strive to improve the coverage of the toolchain platform in the entire process of model and intelligent agent construction, enhance stability, and create a toolchain platform that can support concurrent use by 30% of users across the entire network, supporting the research and development of intelligent agent applications in various provinces. In the future, we will continue to build a toolchain platform ecosystem, promote the integration and collaboration of intelligent agents in different scenarios, achieve cross industry customization, empower thousands of industries, and build applications for the entire industry.

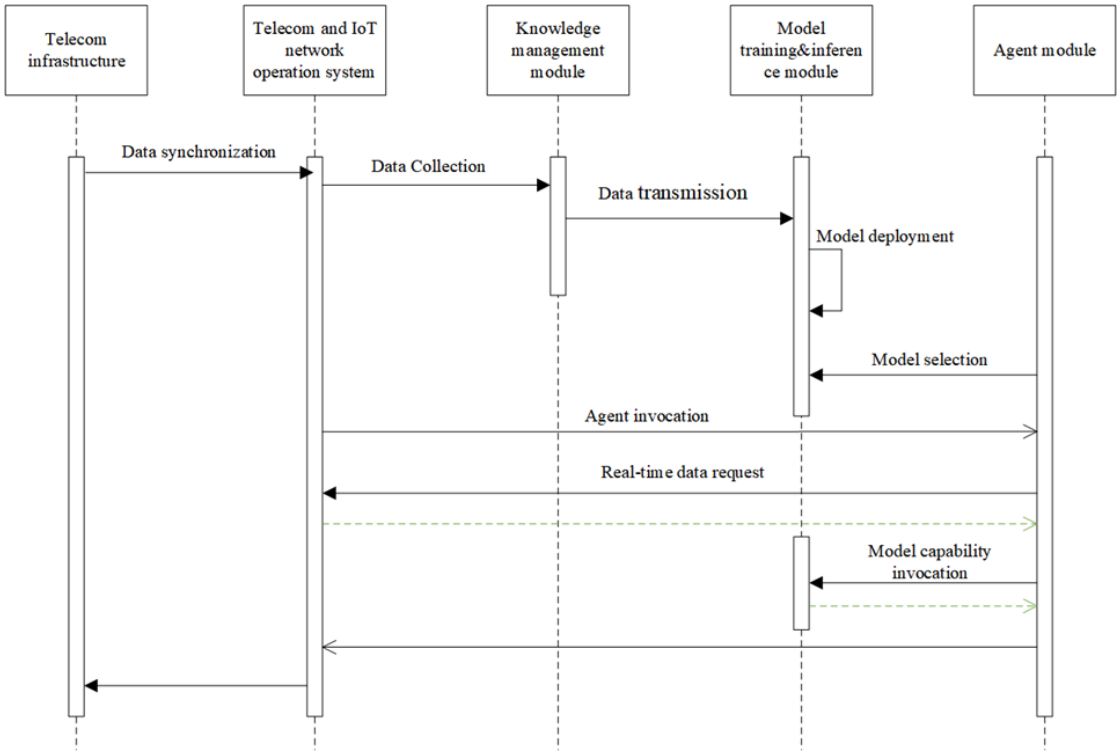
Next, based on this achievement, China Telecom will actively introduce new partners, gradually expand the scale of the alliance, jointly promote the formulation of standards, application for awards, and participation in industry evaluations, and jointly promote the sustainable development and innovation of the network large model industry.

# 3 Use Case Requirements

- **REQ-01**: It is critical that the platform supports large-scale model training and inference through distributed GPU clusters with scheduling, fault tolerance, and mixed-precision optimization.
- **REQ-02**: It is expected that the system enables ingestion and synchronization of multimodal data sources—including structured logs, unstructured text, images, and time-series data—for both real-time and batch processing.
- **REQ-03**: It is expected that the knowledge management module allows users to upload, manage, and preprocess datasets with support for data source integration, permission control, and dataset versioning.

- **REQ-04**: It is critical that the processed datasets are available for training and inference, with selectable training/testing sets and access to a shared compute resource pool for model development and evaluation.
- **REQ-05**: It is expected that the platform provides comprehensive model lifecycle management, including versioning, fine-tuning, deployment, rollback, and real-time monitoring across training, staging, and production environments.
- **REQ-06**: It is expected that the agent module supports construction and orchestration of intelligent agents using tools such as prompt design, model selection, plug-in integration, API linking, and knowledge base access.
- **REQ-07**: It is of added value that users can combine trained models with workflow configurations and knowledge sources to deploy agents tailored to their needs and link them to external business systems.
- **REQ-08**: It is critical that all services, including Model-as-a-Service (MaaS) and agent orchestration, are exposed via secure, scalable, and permission-controlled application programming interfaces (APIs) for external access and automation.

## 4 Sequence Diagram




## 5 References

[1] [ITU_T F.AIAP] Recommendation ITU-T F.AIAP(2025), Framework and requirements for AI agents platform. Available online: https://www.itu.int/md/T25-SG21-250113-TD-WP3-0041/en

[2] [ITU-T F.RF-AIAC-FM] Recommendation ITU-T F.RF-AIAC-FM(2025), Requirements and Framework for AI Agent Collaboration for Foundation Models. Available online: https://www.itu.int/md/T25-SG21-250113-TD-WP3-0054/en

[3] [ITU-T M.3080] Recommendation ITU-T M.3080 (2021), Framework of AI enhanced Telecom Operation and Management (AITOM). Available online: https://www.itu.int/rec/T-REC-M.3080-202102-I

[4] [ITU-T Y.3172] Recommendation ITU-T Y.3172 (2019), Architectural framework for machine learning in future networks including IMT-2020. Available online: https://www.itu.int/rec/T-REC-Y.3172-201906-I

[5] 2023 large model status survey report (issue 1). Available online: https://www.163.com/dy/article/ICVRMR4F05118SRU.html

[6] 2023 Large Language Model Vitality Ranking. Available online: https://mp.weixin.qq.com/s/pXi3O6_XTX3ha0t25heWFA

# Use Case 3: AI-Driven Digital Twins for Energy-Efficient Telco Networks

**Organization**: Tata Elxsi

**Country**: India

**Contact Person(s)**:

Primary contact(s):

**Jesu Maria Antony,** jesumariaantony.s@tataelxsi.co.in

**Shradha Tiwari,** shradha.t@tataelxsi.co.in

**Subhaditya Bhattacharjee** subhaditya.b@tataelxsi.co.in

Secondary contact(s):

**Arun Vignesh,** arun.k5@tataelxsi.co.in

**Kraleti Raja Sekar,** kraletiraja.s@tataelxsi.co.in

**Deepika Muthukumar,** deepika.m3@tataelxsi.co.in

**Bhuvaneshwari S,** bhuvaneshwari.s@tataelxsi.co.in

**Tharani K,** tharani.k@tataelxsi.co.in

**Rajesh M,** rajesh.m2@tataelxsi.co.in

**Korak Bhakta,** korak.b@tataelxsi.co.in

Any other contacts:

**Gomathy**, gomathyp@tataelxsi.co.in

## 1 Use Case Summary Table

| Item | Details |
|---|---|
| Category | 5G |
| Problem Addressed | Telecommunication networks face significant energy inefficiencies due to static provisioning, lack of intelligent automation, and 24/7 infrastructure operation regardless of demand. |

(continued)

| Item | Details |
|---|---|
| Key Aspects of Solution | The solution introduces an AI-native digital twin platform that mirrors the live telecom network in real time, enabling intelligent forecasting of traffic and load conditions. By integrating real-time and historical data, the platform simulates network behaviour and dynamically optimizes configurations to reduce energy consumption without compromising service quality. AI models drive autonomous decision-making, adjusting parameters such as power states, bandwidth, and antenna configurations, while factoring in sustainability goals. The architecture is modular and production-ready, leveraging tools like FastAPI, Kafka, and optional LEXI platform integration for scalability. Designed with green AI principles, the system ensures efficient energy use across both the network and its own AI infrastructure. |
| Technology Keywords | Energy Optimization, Network Intelligence, Digital Twin, Predictive Modelling, AI in Telecom |
| Data Availability | Private (Collected from operational telecom network environments) and Synthetic rule-based data for Training |
| Metadata (Type of Data) | Time-series data, network telemetry, performance logs, configuration data, KPIs |
| Model Training and Fine-Tuning | AI models are trained on historical network behaviour and performance data to predict usage trends, detect anomalies, and recommend energy-saving actions |
| Testbeds or Pilot Deployments | Early-stage pilots in simulated telecom environments with representative traffic and infrastructure |
| Code repositories | NA |

# 2 Use Case Description

## 2.1 Description

In modern telco infrastructure, energy consumption across network elements–from RAN to transport to core - is one of the most critical operational challenges. Despite the integration of more efficient hardware and renewable energy sources, the dynamic and distributed nature of networks results in inefficiencies that are difficult to predict and manage. With 5G, edge computing, and virtualization increasing the complexity, real-time visibility and control over energy usage has become both a necessity and a challenge.

This use case introduces an AI-driven digital twin platform that simulates, monitors, and optimizes energy consumption across the entire telco network. By creating virtual replicas of network assets both passive (cooling, power supplies) and active (base stations, routers, compute nodes) and integrating real-time telemetry from these components, we can build a learning system that predicts energy demand, identifies inefficiencies, and simulates optimization strategies under various network load conditions.

The core of this solution is a set of modular AI/ML models that continuously learn from historical and live data streams to simulate power usage and recommend optimizations. The simulation layer can analyze load patterns, adjust resource allocations, and propose scheduling or

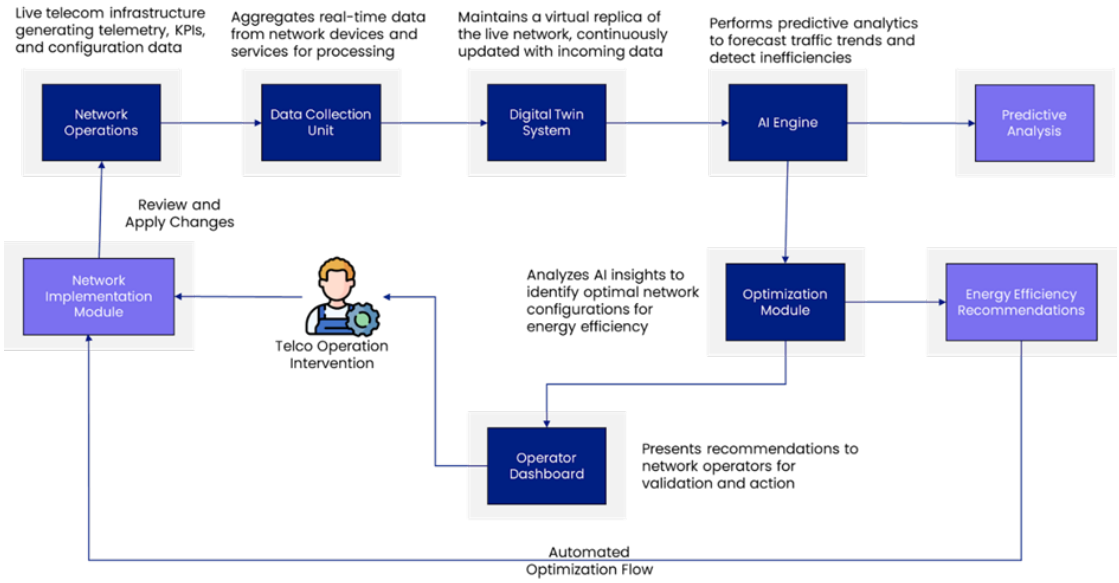
offloading actions that reduce power usage without compromising service quality. Integration with platforms like LEXI (Open source driven) ensures that digital twins are dynamically updated with real-world device data, while connectivity via Kafka and MQTT supports real-time ingestion and orchestration.

Ultimately, this system allows telcos to move beyond static power-saving profiles toward a fully adaptive, context-aware energy management approach. The platform not only improves operational sustainability and aligns with global net-zero goals but also creates tangible cost savings by reducing unnecessary power draw especially during off-peak conditions or under redundant infrastructure states.

## Objectives and Aims:

This use case aims to reduce unnecessary energy usage in telecom operations while maintaining or even improving Quality of Service (QoS). By leveraging AI-based predictions and simulations, the network can become more adaptive and responsive to changing conditions, enabling automated decision-making that balances performance with energy efficiency. The solution also supports predictive maintenance and network self-adaptation, contributing to long-term sustainability objectives and significant operational cost reductions.

Flow Diagram (Functional View)




## Sub-Scenarios:

This solution leverages AI to optimize energy usage through intelligent network management. User admission is dynamically controlled based on real-time load, energy constraints, and QoS requirements, while configurations like antenna tilt and bandwidth are automatically adjusted using predicted usage patterns. The system deactivates idle resources during low-traffic periods to save power and employs traffic load forecasting based on historical data, events, and time-of-day trends. Additionally, energy-aware scheduling shifts non-critical tasks to off-peak hours or times when cleaner energy is available, ensuring efficient and sustainable network operations.

### Impact of Network Digital Twin (NDT):

The Network Digital Twin (NDT) plays a central role by providing real-time synchronization with the actual network, integrating telemetry, configurations, and faults across vendor-agnostic sources. It incorporates AI modules for continuous simulation and feedback. The NDT comprises three key models: a physical topology model representing the network layout, a behavioural model simulating node responses to varying conditions, and a policy model outlining constraints and automation rules. This enables advanced inferences such as predictive failure alerts, energy-saving opportunity mapping, pre/post-configuration impact analysis, and identification of optimal node configurations to reduce energy use without degrading service.

### Network Digital Twin (Data Twin) and AI Inference Relationship:

The Network Digital Twin (or Data Twin) works closely with AI inference engines by continuously feeding them structured network data. Traffic load metrics help predict congestion and guide dynamic resource scaling, while energy consumption logs reveal inefficiencies linked to specific configurations. User behaviour profiles inform hotspot predictions and service quality assessments. KPI trends like latency and throughput help monitor network health and trigger configuration changes, and topological metadata enables AI to apply these decisions intelligently such as offloading or rerouting without violating service-level agreements (SLAs).

### AI-Driven Network Optimization:

AI-driven actions enable automated and intelligent network optimization. These include **dynamic resource scaling** by activating or deactivating radio units and adjusting bandwidth based on demand forecasts. **Energy-efficient configuration updates** are applied, such as tuning power levels or beamforming strategies without compromising service quality. The system also generates **alerts and recommendations** for predicted congestion or failures. AI can **propose topology adjustments** like link rerouting to balance energy and performance. Finally, post-action telemetry feeds into a feedback loop, **continuously refining the AI models** for improved accuracy and adaptability.

### Innovation and Uniqueness:

The solution stands out by offering **dynamic intelligence through adaptive learning** from real-time and historical data, enabling accurate forecasting. It allows telecom operators to safely experiment with **"what-if" scenarios using the digital twin**, fostering proactive decision-making without impacting live operations. By combining deep telecom domain knowledge with AI-driven predictions, it delivers **contextual, interpretable energy optimization**. Its cross-domain applicability supports both fixed and mobile networks, ensuring broad relevance across the telecom landscape.

### Expected Impact:

This solution delivers **significant energy savings** across the telecom network without compromising service quality, using AI-driven predictive models and a real-time digital twin platform. Unlike traditional static or threshold-based energy-saving methods, it offers **proactive network scaling and adaptive configuration** tailored to actual and forecasted demand. Its ability to dynamically deactivate idle resources and shift workloads based on energy availability

or carbon intensity makes it both **intelligent and sustainable**. By integrating high-fidelity simulation, real-time telemetry, and closed-loop automation, the platform ensures **superior network efficiency, cost savings, and measurable support for carbon neutrality goals** all within a **scalable framework** that can adapt to varied network environments and technologies.

**Use Case Status:** In development with prototype demonstrations in test environments

Partners

- Tata Elxsi - Lead implementer and system designer
- Tata Elxsi + Academic Partners - Develop and fine-tune inference models
- Academic Institutions – Support with AI/ML research findings and support model explainability and transparency
- Telecom Operators – Provide operational data and deployment testbeds
- Energy Efficiency Consultants – Validate energy-saving outcomes and benchmarks

## 2.2 Benefits of Use Case

- This use case promotes the adoption of intelligent and sustainable infrastructure by driving innovation in network operations and management. It supports the development of modern industrial systems that are more efficient, adaptable, and forward-looking.
- The system also helps optimize the consumption of energy resources in telecommunications operations, reducing waste and increasing the sustainability of core infrastructure.
- Most importantly, the use case contributes to the reduction of carbon emissions and overall energy consumption in a major industrial sector. This directly supports environmental resilience and climate-conscious operational practices across large-scale telecom deployments.

## 2.3 Future Work

To advance this AI-driven energy optimization platform from concept to wide-scale deployment, the immediate next step involves **scaling pilot testing** within live operator environments. These controlled field deployments will validate the accuracy of energy consumption predictions and the impact of optimization strategies under real-world network conditions. Key metrics such as power savings, QoS maintenance, and system responsiveness will be monitored closely to refine the models.

As part of the system's evolution, we plan to expand the **AI capabilities to support fault prediction and autonomous self-optimization**. This includes enabling the platform to not only forecast potential failures in network components but also to **preemptively adjust configurations** to avoid service degradation, enhancing both energy efficiency and service reliability.

Another major focus area is the automation of response actions, reducing the manual workload for network operators. By integrating **closed-loop control mechanisms**, the system will be able to act on its own recommendations, such as reallocating bandwidth, switching off idle units, or shifting tasks to greener windows based on defined policies and confidence thresholds.

Improving the fidelity of the simulation layer is also crucial. Future iterations will support a broader range of configuration scenarios, including more complex interdependencies across

RAN, transport, and edge layers. Higher simulation accuracy will help operators plan network upgrades and energy strategies with greater confidence.

Strategically, we will **seek partnerships with global telecom operators, standards bodies, and open-source communities** to ensure interoperability and scalability of the platform. This collaborative approach will help align the solution with evolving telco standards and promote broader industry adoption.

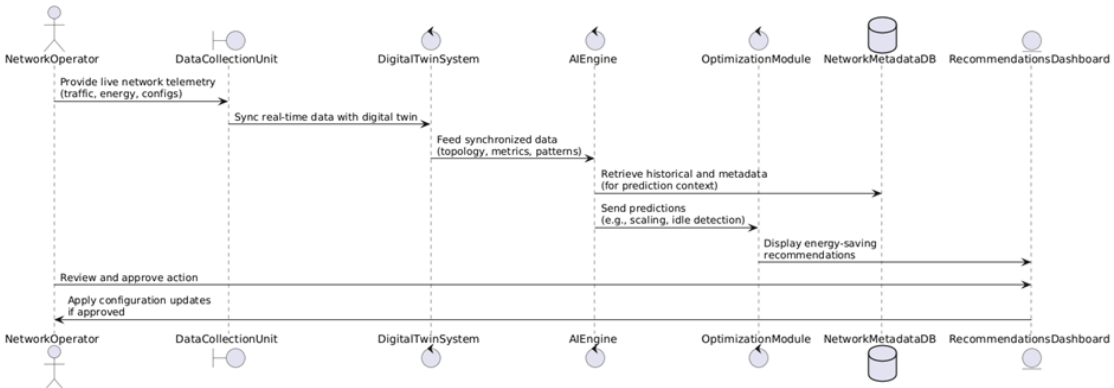
Lastly, a longitudinal study will be initiated to **evaluate the long-term environmental and operational benefits**, quantifying the $CO_2$ reductions and cost savings achieved. This evidence base will support the business case for full-scale rollouts and help shape the next generation of sustainable telco infrastructure.

## 3 Use Case Requirements

- **REQ-01:** It is critical that the system can model a live telecommunications network within a virtualized environment to enable real-time analysis and simulation.
- **REQ-02:** It is expected that the platform ingests real-time operational and performance data to ensure accurate monitoring and adaptive response.
- **REQ-03:** It is critical that artificial intelligence (AI) models are implemented to predict traffic loads, detect anomalies, and identify patterns in energy usage across the network.
- **REQ-04:** It is expected that the system includes mechanisms to simulate multiple network configurations to assess their performance and energy efficiency under varying conditions.
- **REQ-05:** It is critical that the system provides actionable recommendations for optimizing energy consumption based on simulation outcomes and predictive insights.
- **REQ-06:** It is of added value that the platform features a user interface for visualizing the live network state, AI-driven predictions, and system suggestions in an interpretable format.
- **REQ-07:** It is expected that the solution supports deployment in both real-world and simulated telecom environments to ensure flexibility in testing and integration.
- **REQ-08:** It is critical that a closed-loop automation framework is implemented to enable seamless updates to network configurations based on inference results.
- **REQ-09:** It is expected that the system incorporates secure and interpretable feedback loops to validate AI inferences and support ongoing model refinement.

## 4 Sequence Diagram

Please add one line description for the sequence diagram



## 5 References


[1] M. Sanz Rodrigo, D. Rivera, J. I. Moreno, M. Àlvarez-Campana and D. R. López, "Digital Twins for 5G Networks: A Modeling and Deployment Methodology," in *IEEE Access*, vol. 11, pp. 38112-38126, 2023, doi: 10.1109/ACCESS.2023.3267548.

[2] Ezzeddine, Zeinab, Ayman Khalil, Besma Zeddini, and Habiba Hafdallah Ouslimani. 2024. "A Survey on Green Enablers: A Study on the Energy Efficiency of AI-Based 5G Networks" *Sensors* 24, no. 14: 4609. https://doi.org/10.3390/s24144609

[3] UN Sustainable Development Goals (SDG 9: Industry, Innovation and Infrastructure; SDG 13: Climate Action), [Online]. Available: https://sdgs.un.org/goals.

[4] ITU/UN, "AI for Good Global Summit – AI and Digital Twins for Sustainability in Telecommunications," [Online]. Available: https://aiforgood.itu.int.

[5] P. Singh and S. Tiwari, "Digital Twins and Artificial Intelligence," ResearchGate, 2024. [Online]. Available: https://www.researchgate.net/publication/386431226_Digital_Twins_and_Artificial_Intelligence.

[6] ITU-T, *Framework for Network Digital Twin*, Recommendation ITU-T Y.3090, Feb. 2022. [Online]. Available: https://www.itu.int/rec/T-REC-Y.3090-202202-I!!PDF-E

[7] H. X. Nguyen, R. Trestian, D. To, and M. Tatipamula, "Digital twin for 5G and beyond," IEEE Commun. Mag., vol. 59, no. 2, pp. 10–15, Feb. 2021, DOI: 10.1109/MCOM.001.2000343.

[8] C. Zhou, H. Yang, X. Duan, D. Lopez, A. Pastor, Q. Wu, M. Boucadair, and C. Jacquenet ( Oct. 2022 ). Digital Twin Network: Concepts and Reference Architecture. Accessed: Dec. 1, 2022. [Online]. Available: https://datatracker.ietf.org/doc/draft-irtf-nmrg-network-digital-twin-arch

# Use Case 4: Integrated AI and Wireless for Sustainable Mobile Network Evolutions

**Organization:** Shanghai University

**Country**: China

**Contact Person(s)**:

Zhiyuan Jiang, jiangzhiyuan@shu.edu.cn

Yi Shi, yishi1996@shu.edu.cn

Sheng Zhou, sheng.zhou@tsinghua.edu.cn

## 1 Use Case Summary Table

| Item | Details |
|---|---|
| Category | 5G |
| Problem Addressed | The dominance among a few vendors in the mobile networks industry leads to increased costs and restricts innovation [1]. These closed ecosystems hinder transparency and coordination efforts, and the lack of open-source designs widens the gap between theoretical advancements and real-world implementation. |
| Key Aspects of Solution | The three modules for solving this problem are Echo for Open-Sourcing a library for 5G, 4G, GNSS, LoRa, etc, Zoozve for creating an integrated toolchain for AI and wireless baseband processing (WBP), and Venus for creating a Multi-Core Dataflow-Driven RISC-V Domain Specific Architecture for Integrated AI and WBP. The project aims to promote open-source baseband architecture, which encourages innovation and allows more companies to enter the market. By integrating AI with wireless baseband processing, it accelerates technology iteration and fosters industrial development. The project reduces reliance on proprietary hardware and emphasizes open protocols, enabling small and medium enterprises and innovators to compete in telecommunications. Additionally, the RISC-V architecture, optimized for AI-WBP synergy, enhances energy efficiency and promotes sustainable industry practices. |
| Technology Keywords | Wireless Baseband Processing, RISC-V, Wireless Communication, RAN (Radio Access Network), Chip Architecture. |
| Data Availability | Not applicable |
| Metadata (Type of Data) | Not applicable |

(continued)

| Item | Details |
|---|---|
| Model Training and Fine-Tuning | This architecture enables AI and wireless baseband integration. |
| Testbeds or Pilot Deployments | The Echo Testbench serves as an open-source wireless testbed providing a full-stack development environment [4] |
| Code repositories | The source codes for the Echo computing architecture, Venus LLVM tool-kit for the construction of highly optimized compilers, optimizers, and run-time environments, and Zoozve, a compiler, are open-sourced [2][3][4]. |

# 2 Use Case Description

## 2.1 Description

### Background and Motivation

The mobile communications industry faces a sustainability crisis. While 5G delivers performance improvements, it has failed to catalyze disruptive applications, leaving operators burdened with exorbitant deployment costs. Industry leaders have even voiced skepticism about 6G. This dilemma stems fundamentally from the closed nature of traditional baseband architectures: hardware relies on rigid ASICs or DSPs, and software ecosystems remain proprietary, creating three critical contradictions—limited hardware choices stifle cross-domain innovation, the absence of open-source platforms impedes academia-industry collaboration, and protracted development cycles cannot keep pace with 6G's rapid evolution. The essence of 6G lies in transforming AI from a "service carried by networks" to a "native network capability," manifesting as dual paradigms: AI4NET (end-to-end AI optimization of air interfaces and operations) and NET4AI (native integration of communication-computing-sensing capabilities within network elements). This convergence demands architectural reinvention that supports large-scale distributed AI applications while embedding trustworthiness—encompassing security, privacy, and resilience—as foundational attributes. The establishment of the AI-RAN Alliance **[5]** in February 2024 signals industry momentum toward GPU-accelerated AI-RAN integration, though its high power consumption and terminal deployment limitations remain unresolved.

### Technological Landscape and Innovation Gap

(1) Open-source projects like OAI/srsRAN rely on x86/CPUs, delivering subpar performance without AI-native support;

(2) AI-RAN leverages GPUs for joint AI-RAN deployment but suffers from communication energy efficiency far inferior to ASICs and cannot be deployed on end-user devices;

(3) Modem + NPU discrete solutions incur high inter-chip latency and lack flexibility for academic research. This closed ecosystem creates two critical fractures: cutting-edge academic research struggles to translate into industrial solutions, while operators remain vendor-locked with diminished architectural flexibility.

These challenges underscore the imperative for an open-source, fully software-defined baseband architecture with deep AI-communication convergence—the core mission of the Venus initiative.

## Our Solution

The Venus initiative delivers a comprehensive ecosystem incorporating silicon (Venus), toolchains (Zoozve), and software stacks (Echo). The Venus Chip Architecture **[6]**, implemented in 40nm technology, pioneers the world's first RISC-V domain-specific architecture that fuses AI and communications. Innovative features include a cacheless manycore architecture eliminating energy waste through dataflow-driven computing, optimized for wireless baseband processing (WBP); hierarchical dataflow scheduling that abstracts WBP into Directed Acyclic Graphs, using a "Pack-n-Ship" mechanism to reduce data transfer overhead by 60%; and heterogeneous configurable clusters supporting dynamic core combinations, achieving 288Mbps link throughput on SMIC 40nm prototypes with FFT/BP algorithm performance reaching 2.3x/2x that of commercial hardware. Additionally, the Zoozve Compiler introduces the first "Strip-Mining-Free" RISC-V vector extension compiler. It features an asymmetric instruction set, eliminating strip-mining overhead for ultra-long vector operations, reducing dynamic instruction counts by 344x/76x for FFT/DOT; arbitrary register grouping to overcome fixed-register constraints in RVV, optimizing AI operations like 2D convolution; and a heterogeneous compilation framework that unifies CPU/NPU memory management, leading to 7.06x higher code density and 5.58x memory efficiency in CNN models. Lastly, the Echo Testbench serves as an open-source wireless testbed providing a full-stack development environment. It includes a high-performance operator library covering 5G/4G/GNSS/LoRa and AI operators, comparable to Intel AVX/ARM Neon; a three-layer programming model consisting of a task layer (modules in Venus ISA), a DAG layer (network assembly via .bas files), and a scheduler layer (L1 hardware control); and full-stack emulation that enables closed-loop development, significantly reducing traditional baseband R&D costs.

## Validation and Application Scenarios

FPGA prototypes have demonstrated viability across three critical scenarios: RedCap Lightweight Terminals have successfully completed commercial base station communication tests, validating architectural support for mid-low-speed IoT; high-precision GNSS positioning achieves low-power, full-band reception for autonomous vehicles and drones; and AI-enhanced communications develop real-time AI ionospheric delay prediction algorithms, showcasing communication-sensing-intelligence fusion. The testbed has open-sourced all LTE/NR PHY modules and foundational AI operators, with full protocol stack implementation forthcoming.

## Conclusion

The Venus project represents a paradigm shift in baseband processing. Technically, it achieves the first ISA-level fusion of communications and AI on RISC-V DSA, overcoming the "impossible trinity" of performance, efficiency, and elasticity through dataflow-driven manycore architecture. Industrially, it constructs the first comprehensive open-source baseband stack, dismantling closed ecosystems and providing operators with vendor-agnostic foundations. Ecologically, the Echo testbed acts as a "PyTorch for wireless communications," accelerating the translation between academia and industry. When questioned about the necessity of 6G, Venus responds by emphasizing that the essence of 6G lies not in higher frequencies or more antennas, but in transforming networks into intelligent entities through open architecture, laying the foundation for future innovations.

### Partners

- **Tsinghua University** [7], **Deep Transcend Co., Ltd.**

## 2.2 Benefits of use case

- The project promotes open-source, software-defined baseband architecture, encouraging innovation and allowing more companies to enter the market. By integrating AI with Wireless Baseband Processing (WBP), the project aims to accelerate technology iteration, fostering industrial development.
- By reducing reliance on proprietary hardware and emphasizing open protocols, the project allows small and medium enterprises (SMEs) and innovators to participate, levelling the playing field in the telecommunications industry.
- The use of a RISC-V architecture optimized for AI-WBP synergy enhances energy efficiency, contributing to more sustainable telecom infrastructure. The emphasis on open-source and standardized designs encourages responsible production practices and sustainable industry transformation.

## 2.3 Future Work

Future work on the open-source baseband architecture will focus on expanding protocol support, advancing AI integration, and enhancing energy-efficient design. Efforts will be made to establish robust collaborative platforms that link academia and industry to streamline innovation. Opportunities to partner with leading telecommunications research bodies will be explored to refine the architecture and address emerging industry needs. These initiatives aim to develop a versatile platform that industry leaders, researchers, and policymakers can access globally. Ultimately, the goal is to empower stakeholders with cutting-edge tools for strategic decision-making, thereby fostering sustainable industry growth and innovation.

The roadmap focuses on three key areas: technical advancement, ecosystem development, and scenario expansion. Technical advancements aim to expand protocols towards 6G, enhancing AI-native capabilities such as federated learning frameworks, and achieving breakthroughs in energy efficiency with near-memory computing. Ecosystem development will be highlighted at ITU AI for Good, with the launch of expanded operator libraries and multi-functional toolchains to create academia-industry collaboration platforms similar to PyTorch's ecosystem. Scenario expansion targets emerging domains like aerial economy and communications in extreme environments, transforming basebands from "closed black boxes" to "open innovation platforms." This approach addresses 6G's sustainability challenges by unifying performance, flexibility, and economy through an open DSA architecture.

# 3 Use Case Requirements

- **REQ-01**: Open Protocol Implementation

  It is critical that open protocols are developed and implemented to reduce reliance on proprietary hardware and enhance interoperability.

- **REQ-02**: AI Integration with WBP

  It is expected that AI capabilities will be integrated with Wireless Baseband Processing to accelerate technology iteration and improve performance.

- **REQ-03**: Energy-Efficient Architecture

It is critical that a RISC-V domain-specific architecture is designed, focusing on energy efficiency and adaptability to future communication standards.

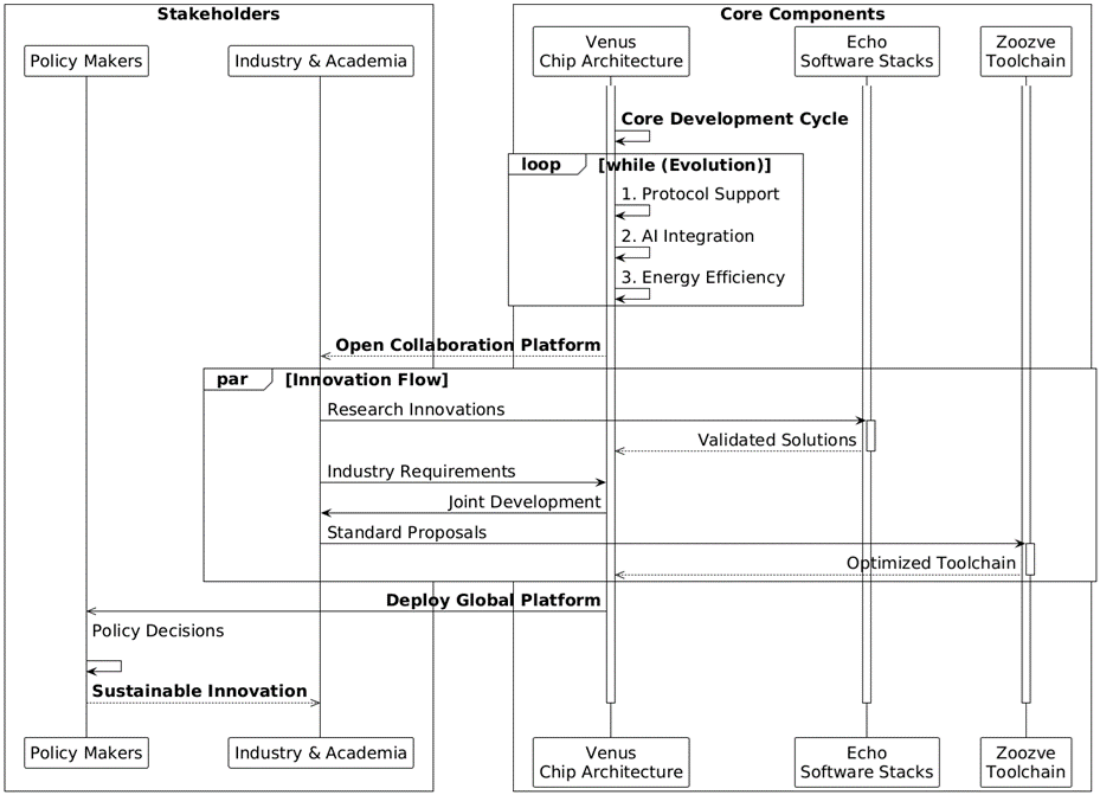
- **REQ-04**: Collaborative Platforms

  It is of added value that open collaborative platforms are established to bridge the gap between academia and industry, facilitating rapid translation of research into practice.

- **REQ-05**: User-Friendly Toolchain

  It is expected that a user-friendly toolchain will be created for seamless deployment of algorithms into hardware, supporting dynamic resource sharing and scheduling.

## 4 Sequence Diagram




## 5 References

[1] Startech Networks, "Unleashing Innovation: The Rise of Open RAN in Telecommunications," StartechNetworks.com, Mar. 25, 2024. [Online]. Available: https://startechnetworks.com/unleashing-innovation-the-rise-of-open-ran-in-telecommunications

[2] ACELab-SHU, "Zoozve: A strip-mining-free RISC-V vector extension with arbitrary register grouping," GitHub repository, 2025. [Online]. Available: https://github.com/ACELab-SHU/LCTES-25-Artifact

[3] ACELab-SHU, "Venus-Compiler," GitHub repository, 2024. [Online]. Available: https://github.com/ACELab-SHU/Venus-Compiler.

[4] ACELab-SHU, "ACE-Echo," GitHub repository, 2024. [Online]. Available: https://github.com/ACELab-SHU/ACE-Echo

[5] AI-RAN Alliance, "AI-RAN: Accelerating AI-native transformation in future wireless networks," [Online]. Available: https://ai-ran.org/

[6] L. Jiang *et al*., "A Hierarchical Dataflow-Driven Heterogeneous Architecture for Wireless Baseband Processing," 2025.[Online]. Available: https://arxiv.org/pdf/2402.18070

[7] Tsinghua University, "General Information of Tsinghua University," [Online]. Available: https://www.tsinghua.edu.cn/

# Use Case 5: Generative Artificial Intelligence Content Security Detection and Protection Platform

**Organization**: China Telecommunications Corporation

**Country**: China

**Contact Person(s)**:

Primary contact:

Hongdan Ren, renhd@chinatelecom.cn

## 1 Use Case Summary Table

| Item | Details |
|---|---|
| Category | 5G |
| Problem Addressed | The increasing use of AI-generated content presents challenges in detecting and safeguarding against harmful or inappropriate material |
| Key Aspects of Solution | • High-quality AI security corpus pipeline based on cognitive generation<br>• Continuous output of evaluation data and training corpora<br>• Automated evaluation and protection of algorithm models |
| Technology Keywords | Generative AI, Content Security, Detection, Protection, AI Security Corpus, Distillation Architecture, Semantic Sensitivity, |
| Data Availability | Not Available |
| Metadata (Type of Data) | Text, audio, visual |
| Model Training and Fine-Tuning | Developed a new generation of ideological security assessment large model to address the content security challenges in the field of generative artificial intelligence. By constructing a million-level political security corpus and adopting knowledge distillation techniques to transfer the reasoning capabilities of the deep thinking model, combined with instruction fine-tuning (SFT) to achieve domain adaptive training, the model's accuracy in identifying subtle ideological expressions has been significantly enhanced. Both the false positive rate and the false negative rate are approximately 6%. |
| Testbeds or Pilot Deployments | Not Available |
| Code repositories | Not Available |

# 2 Use Case Description

## 2.1 Description

The platform is designed to address the growing challenges of content security detection and protection in the context of AI-generated content, with the goal of ensuring that AI technologies are used responsibly and for the public good. As large model technologies rapidly evolve, they introduce increased risks related to the generation and dissemination of harmful or inappropriate content. Traditional content security methods rely heavily on manual review, which is both costly and inefficient at scale.

To overcome these limitations, the platform integrates a high-quality AI security corpus pipeline built on cognitive generation techniques. It enables the continuous production of evaluation data and training corpora. Furthermore, the platform offers automated evaluation and protection mechanisms for algorithmic models.

Technological Breakthroughs and Innovations:

1) Deep Thinking Distillation Architecture

   A layered knowledge distillation framework is designed to extract deep reasoning capabilities from general large models, specifically tailored for ideological risk assessment **[2]**. A dual-channel feature fusion mechanism is constructed to enable the collaborative optimization of semantic understanding and security assessment **[1]**. This approach ensures that both linguistic context and potential security implications are jointly evaluated, resulting in more accurate and reliable risk detection.

2. Semantic Sensitivity Perception Instruction System

   A multi-granularity sensitive instruction set is innovatively designed, encompassing five major evaluation dimensions, such as politics and business, to achieve adaptive feature extraction at different sensitivity levels.

3. Scene Adaptive Review Engine

   A configurable review strategy matrix is designed to enable multi-mode switching between strict control and flexible adaptation.

Deployment Status:

Deployed across the entire China Telecom network, covering 622 million Chinese users

## 2.2 Benefits of the use case

The platform promotes the healthy development of AI technology by providing secure and reliable AI applications for industrial innovation and infrastructure development. It assists regulatory bodies in effectively supervising AI applications, maintaining social order and cybersecurity. The platform fosters collaboration and exchange in the field of AI security, collectively advancing the sustainable development of AI technology.

## 2.3 Future Work

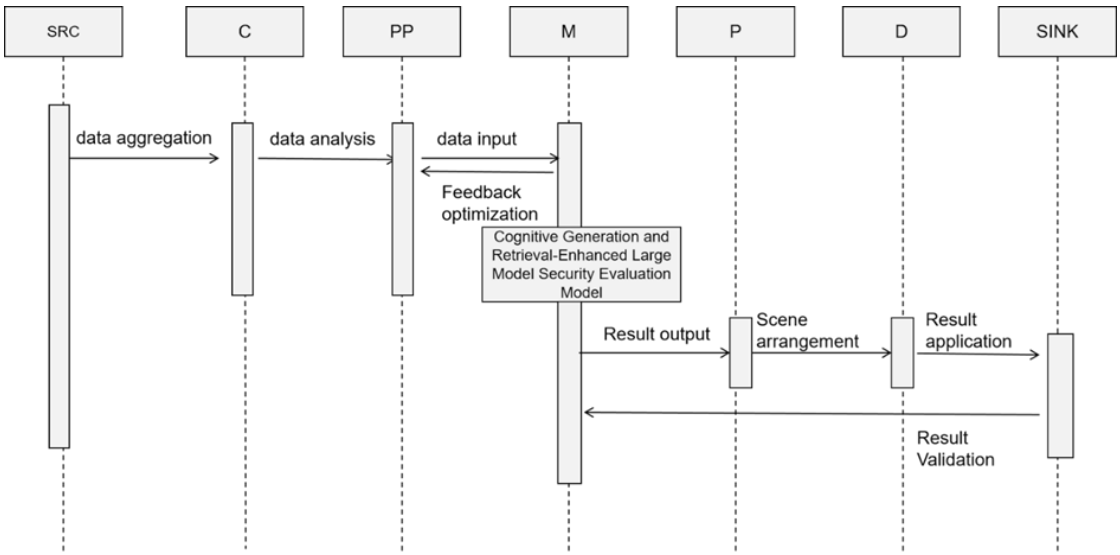
- Refine the ideological assessment model to enhance its accuracy and efficiency, and develop additional types of detection models to cover a wider range of risk scenarios.

- Create new AIGC content security detection and protection products tailored to various scenarios, such as social media, news media, e-commerce platforms, and other vertical industries **[3]**.
- Participate in the development of standards and regulations related to AIGC content security detection and protection, promoting the healthy development of the industry.
- Establish reference tools and simulation environments for AIGC content security detection and protection to help users better understand and apply the platform.

## 3 Use Case Requirements

- **REQ-01**: It is critical that a high-quality AI security corpus pipeline based on cognitive generation is both deployed, as it forms the foundation for training, evaluating, and improving the system's ability to detect and respond to security risks in AI-generated content.
- **REQ-02**: It is critical that the system is capable of continuously generating evaluation data and training corpora.
- **REQ-03**: It is critical that algorithmic models are evaluated and protected automatically to ensure the trustworthiness of AI systems in operational environments.

## 4 Sequence Diagram




- SRC (source): This node is the source of data that can be used as input to the ML pipeline.
- C (collector): This node is responsible for collecting data from one or more SRC nodes.
- PP (preprocessor): This node is responsible for cleaning data, aggregating data or performing any other preprocessing needed.
- M (model): This is a machine learning model
- P (policy): This node enables the application of policies to the output of the model node.
- D (distributor): This node is responsible for identifying the SINK(s) and distributing the output of the M node.
- SINK: This node is the target of the ML output on which it takes action.

## 5 References

[1] J. Kim, M. Hyun, I. Chung, and N. Kwak, "Feature Fusion for Online Mutual Knowledge Distillation," arXiv preprint:1904.09058*, Apr. 2019. [Online]. Available: https://arxiv.org/pdf/1904.09058

[2] Y.-C. Chen, Z. Gan, Y. Cheng, J. Liu, and J. Liu, "Distilling the Knowledge of BERT for Text Generation," *arXiv preprint:1911.03829*, Nov. 2019. [Online]. Available: https://arxiv.org/pdf/1911.03829

[3] K. Wang, Z. Fu, L. Zhou, and Y. Zhu, "Content Moderation in Social Media: The Characteristics, Degree, and Efficiency of User Engagement," in *Proc. 2022 3rd Asia Symposium on Signal Processing (ASSP)*, Bangkok, Thailand, Dec. 2022, pp. 86–91, doi: 10.1109/ASSP57481.2022.00022

# Use Case 6: The "CloudTop": AI-driven Marketing and Service Intelligentized Applications of China Telecom Cloud

**Organization:** China Telecom Cloud Technology Co., Ltd

**Country**: China

**Contact Person(s)**:

Aijiao Wang, wangaijiao@chinatelecom.cn

Qifang Liu, liuqf2@chinatelecom.cn

Wenwei Li, liww9@chinatelecom.cn

## 1 Use Case Summary Table

| Item | Details |
|---|---|
| Category | 5G |
| Problem Addressed | There is an excessive amount of low-level, repetitive work in simple scenarios, and some of these tasks have a negative impact on employees' physical and mental health, as well as their career development. The Conventional digitalization solutions often struggle to support the complex needs of China's industries in their digitalization and intelligence transformation. |
| Key Aspects of Solution | Develop a vertical domain model tailored to the enterprise by integrating external enterprise knowledge into a general-purpose large language model. Implement a Retrieval-Augmented Generation (RAG) system that leverages high-quality external knowledge bases to enhance the accuracy and reliability of responses. Utilize Automatic Speech Recognition (ASR) models to convert speech into text, enabling precise recognition of customer intent. Apply efficient semantic retrieval techniques, such as dense vector retrieval, and integrate multiple models to balance speed and accuracy for tasks including question summarization, sentiment analysis, entity extraction, and product categorization. Additionally, build self-reflective agents capable of generating self-feedback to optimize their decision-making logic over time. |
| Technology Keywords | Natural Language Processing (NLP), Large Language Models (LLMs), Text Classification, Data Analytics, Fine-tuning, Retrieval-Augmented Generation (RAG), Prompt Engineering, ASR, Sentiment Analysis, Self-reflection Agent |
| Data Availability | The project uses both public and private data [2]. The knowledge documents part of the model training data are accessible from the China Telecom Cloud official documentation webpage [8]. |
| Metadata (Type of Data) | Audio and voice data, text data, image data, numerical data, and label data are generated during customer service interactions and enterprise solution delivery. |

(continued)

| Item | Details |
|---|---|
| Model Training and Fine-Tuning | Models: ASR is powered by China Telecom Cloud's self-developed ASR technology. NLP uses TeleChat2 and Huize (self-developed large language models by China Telecom), DeepSeek, and BGE.<br>Training and Fine-Tuning: The ASR model is trained on annotated customer voice data. NLP models are fine-tuned with text, image, and numerical data to develop a vertical domain model capable of answering cloud industry-specific queries. By incorporating annotated data for specific tasks, we guide the model to optimize for those tasks. For example, we train a classification model for customer service scenarios using large amounts of labeled classification data, which helps Customer Service Representatives automatically tag product categories and issue categories for customer service requests. |
| Testbeds or Pilot Deployments | The use case is part of a larger product development; It's now in production at China Telecom Cloud. Services have been deployed through the Ctyun portal [7] |
| Code Repositories | Not Available |

# 2 Use Case Description

## 2.1 Description

Traditional marketing and service models are labour-intensive, with a lot of repetitive, low-skilled work in simple scenarios. This leads to high turnover rates and doesn't support employees' well-being or long-term career development. At the same time, one-size-fits-all digital solutions often fall short in supporting industry-wide digital and intelligent transformation.

To address the above industry challenges, this case proposes an advanced solution. Based on leading frontier AI technologies such as NLP, LLM, Artificial Intelligence Generated Content (AIGC), RAG, Base General Embedding (BGE), Prompt Engineering, ASR, and Sentiment Analysis, it upgrades the technologies and working modes for tasks like intent understanding, knowledge retrieval, intelligent question answering, solution recommendation, and risk monitoring. It has realized intelligent marketing service application capabilities such as intelligent robots, intelligent knowledge bases, intelligent recommendations, intelligent assistance, and risk monitoring. It is committed to creating emerging technology positions, enhancing labour productivity and employee satisfaction, providing demand insights and customized solutions, promoting rational consumption of digital services, accelerating the digital transformation of enterprises, and achieving high-level digital economic growth.

Consider the scenario of using intelligent assistance to reduce the low-level repetitive work of service personnel. This case fully exploits the value of diverse data. By using text data, image data, and numerical data to conduct fine-grained fine-tuning of the model, it helps the model complete the training process, turning it into a vertical domain model that is deeply adapted to the needs of specific industries and can accurately recommend knowledge in professional fields.

During the actual service process carried out by service personnel, the system uses ASR technology to convert customers' voice data into text content and extract questions.

Then, it initiates retrieval from the RAG (Retrieval-Augmented Generation) built on vertical domain documents, and the retrieved content and extracted questions are fed into the large language model, which generates precise and detailed recommendations according to prompts, effectively improving the efficiency of marketing services and enhancing both customer satisfaction and employees' labour productivity **[1]**. To further enhance the Retrieval-Augmented capability, the RAG capability in this case adopts deep semantic encoding based on the BGE (Embedding Model) to improve the generalization ability of knowledge retrieval. It is pre-trained on a million-level exclusive dataset and dual-reinforced by supervised fine-tuning and trainer feedback to train a specialized vector model for vertical domains. Through vectorized retrieval, it achieves rapid data positioning and precise recall. It utilizes the LLM and Prompt Engineering to optimize the generation strategy, improving the accuracy of intelligent recommendations **[3]**.

Currently, the RAG is built on a modular technology stack based on the LangChain framework. Through the unified service interfaces of RESTful API and SDK, it has achieved standardized implementation throughout the entire process, constructed a reusable enterprise-level solution, and can be quickly replicated in other industries.

The self-reflection agent is applied in two primary scenarios, intelligent assistance and intelligent quality inspection. In the context of intelligent assistance, it plays a critical role in refining the output of a question summarization model. After analyzing the conversation between a customer and a customer service representative, the model generates an initial summary of the issue. This summary is then evaluated by the self-reflection agent, which uses the original conversation as a benchmark. The agent assesses the summary based on criteria such as accuracy, comprehensiveness, objectivity, and logical reasoning. If the agent assigns a score below the predefined threshold, it identifies any deficiencies in the content and generates targeted improvement suggestions. These suggestions, such as highlighting missing information, are formulated into new prompts for the summarization model. The model then regenerates the summary using this refined guidance. This evaluation and revision process is repeated until the summary meets the quality threshold, ensuring that the final output is both accurate and complete.

## 2.2 Benefits of the use case

In terms of “Decent Work”: The organization-wide adoption of AI not only helps service and marketing personnel master new skills to adapt to the AI-infrastructure-driven employment environment, reducing employment inequality among young people caused by the lack of new technology training, but also creates emerging jobs such as AI data labeling. Meanwhile, agents independently extract knowledge and solutions, complete quality inspection evaluations, and update quality inspection report data, increasing the coverage rate of knowledge compilation and quality inspection from 10% (manual) to 100%. Human resources from original positions have transformed into AI annotators, assisting in providing feedback to the "intelligent evolution engine" of meta-review agents, enabling continuous learning and self-evolution of AI capabilities.

By integrating customer clustering analytics and historical behavioral data to build dynamic customer profiles, our AI system combines semantic recognition and query reformulation to accurately identify customer needs. Leveraging the Transformer architecture, it instantly matches the most relevant knowledge documents or scenario-based solutions, delivering

them directly to service or sales teams. For example, in customer service, our "AI-assisted + human-in-the-loop" model uses ASR for real-time transcription, natural language processing for intent extraction, and retrieval-augmented generation for instant answers. AIGC generates summaries, together improving service efficiency by 24% and boosting marketing by 15%. The 24/7 chatbot reduces night shifts by 30%, improving working conditions for employees.

In terms of "Economic Growth,": CloudTop leverages Transformer-based models for deep intent understanding and tailored solutions, enabling efficient, low-cost digital transformation. This helps large firms improve productivity and SMEs accelerate innovation, driving high-quality digital growth. Through cluster analysis of customers' internal behavioral information and external characteristic information in the enterprise, different risk scenarios are preset according to the usage of security products. When customers initiate consultations, we will accurately identify whether they need security protection by combining customer intent, clustering results, and risk scenarios. At the same time, customized security reinforcement and highly available security solutions are provided based on customers' resource usage to support the sustainable development of customer businesses and promote healthy economic growth.

Its standardized APIs and container-based deployment make integration quick and easy, allowing all industries to scale AI-powered marketing and services, enhance staff productivity, and increase business value.

## 2.3 Future Work

Future work on this case will focus on model development, creating new variations/extensions for the same use case, and standards development related to the use case. We plan to expand our case's influence in two key areas, application optimization and social contribution.

Application optimization:

The resources from the awards will be allocated to support the ongoing development, platform construction, and promotion of the CloudTop. Expand the application scope of Agents from post-event knowledge extraction and quality inspection to mid-event and pre-event stages, establish the capability to predict customer behavior, provide prompt capabilities for correct behaviors and interception capabilities for abnormal behaviors for service and marketing personnel, and eliminate penalties for employee errors. We will do further research on semantic understanding, sentiment analysis, and multi-round dialogue technologies to enhance the interactive experience and problem-solving efficiency of robots. We will also optimize machine learning algorithms to improve the accuracy and response speed of intelligent agents, and enhance reasoning capabilities. Additionally, the dataset will be improved continuously, and a large-scale high-quality corpus and a high-level data annotation mechanism will be built to promote model iteration and upgrading. Based on the high-quality customer service and marketing dataset precipitated from deep applications, Agents refine the dataset and open it to the public, promoting collaborative innovation across industries.Meanwhile, we will keep a close eye on regulatory requirements and the forefront of industrial technologies, and continuously strengthen data security and privacy protection.

Social contribution:

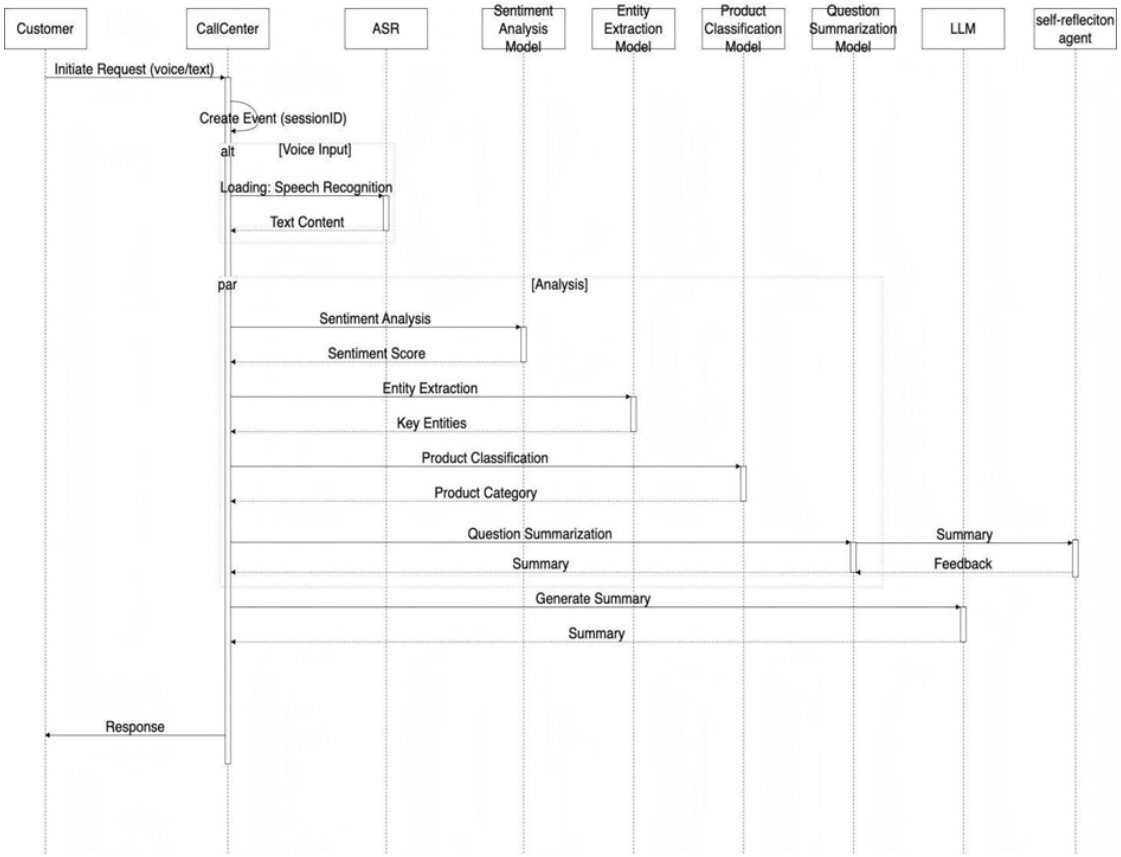
The resources will be utilized to strengthen cooperation with the International Telecommunication Union (ITU) and contribute to the formulation of relevant ITU standards. Based on the deep

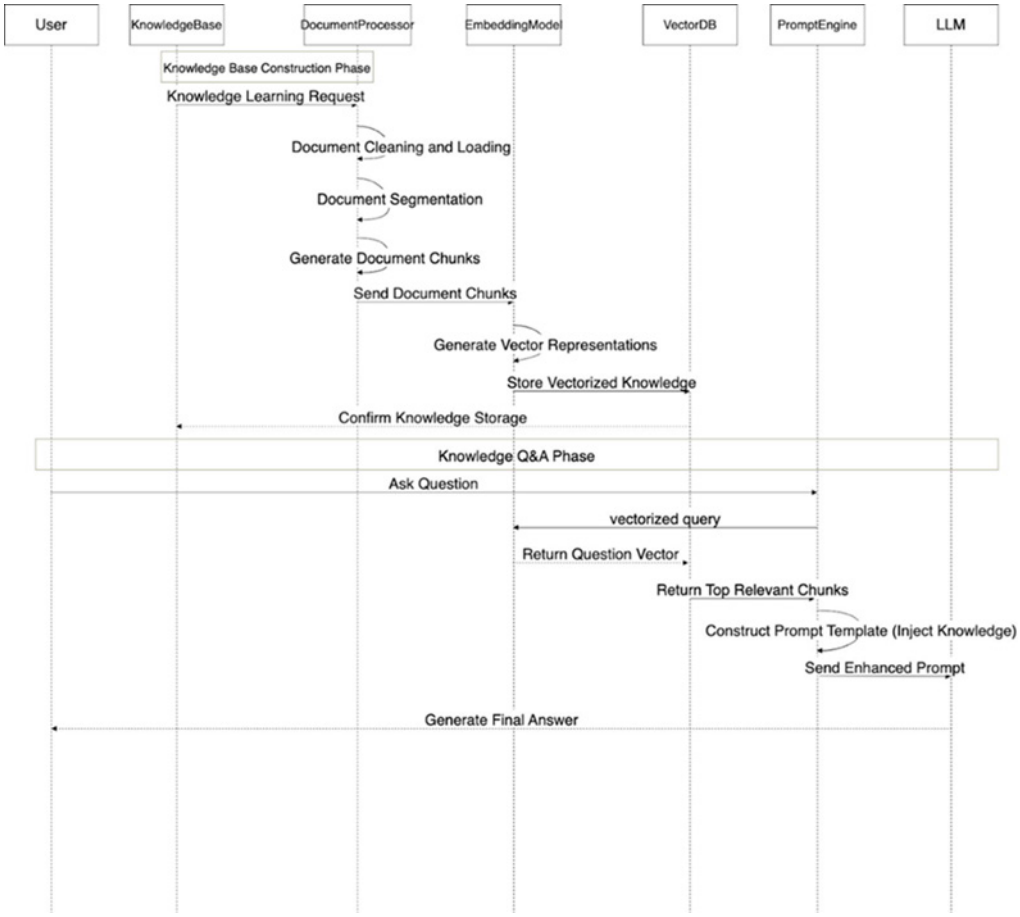
application and accumulation of experience in production scenarios, we will jointly develop Agent technical standards for the customer service and marketing field with the China Academy of Information and Communications Technology (CAICT) and the International Telecommunication Union (ITU), helping practitioners worldwide in the service and marketing industry standardize their work and achieve decent labor. By enhancing collaboration with industry partners, we aim to jointly propel the development and application of technology, fostering a vibrant and innovative ecosystem. Drawing on the practical experiences and professional expertise of the "CloudTop" AI-enabled marketing service applications, we will disseminate this knowledge to more countries and industries in need. Adhering to the principles of sustainable development, we will further explore viable pathways for artificial intelligence technology across different nations and sectors, striving to promote the goals of decent work and economic growth, and contributing to the construction of a more equal, equitable, and sustainably developed future.

## 3 Use Case Requirements

- **REQ-01**: It is critical that cross-functional teams of service marketers, AI operators, and developers be established to continuously enhance the model capabilities and drive the expansion of application scenarios.
- **REQ-02**: It is critical that a full-process digital management system be available to achieve the seamless integration of customer service request data (such as consultations, fault reports, complaints, suggestions, satisfaction evaluations, etc.) and various solution data.
- **REQ-03**: It is critical that a friendly, interactive interface be provided, enabling both employees and customers to use it conveniently and improving the service efficiency and experience.
- **REQ-04**: It is critical that a precise response control system for sensitive topics such as politics and religion be established, and that for the parts involving data privacy and security, sensitive information be ensured to be handled in a responsible manner.
- **REQ-05**: It is critical that its capability boundaries be clearly stated, and in case of situations that cannot be handled or are uncertain, its limitations be proactively declared.

# 4 Sequence Diagram






## 5 References

[1] G. Zhao, Y. Tu, J. Yang, S. Zheng, and Q. Zhang, "Design and implementation of a large model RAG enrollment Q&A system based on domain data enhancement," *DOI:* 10.3969/j.issn.1003-6970.2025.01.052.

[2] Y. Ju, W. Wang, Y. Zhang, et al., "CMQA: A dataset of conditional question answering with multiple-span answers," in *Proceedings of the International Committee on Computational Linguistics*, Gyeongju, Republic of Korea, 2022.

[3] P. Lewis, E. Perez, A. Piktus, et al., "Retrieval-augmented generation for knowledge-intensive NLP tasks," *Advances in Neural Information Processing Systems*, vol. 33, pp. 9459–9474, 2020.

[4] CTYUN, "Intelligent Robots", [Online]. Available: https://ocss.ctyun.cn/ucfront/uc_channel_access/webchat/pages/robot.html

[5] "China Telecom Cloud won multiple awards at the 23rd China IT User Satisfaction Conference", [Online]. Available: https://mp.weixin.qq.com/s/_DqndMzI7Pb_5tG7WR0Clg

[6] "The golden words from China Telecom Cloud experts at the 2024 Trusted Cloud Conference", [Online]. Available: https://mp.weixin.qq.com/s/Z_7VUM7TcNlyD4ILUQwl5g

[7] China Telecom Cloud, "Robot Chat Interface," [Online]. Available: https://ocss.ctyun.cn/ucfront/uc_channel_access/webchat/pages/robot.html

[8] China Telecom Cloud, "产品文档 – 帮助中心," [Online]. Available: https://www.ctyun.cn/document/

# Use Case 7: "Di Ting" Customer Service Agent

**Organization**: China Telecom Co., Ltd. Hubei Branch

**Country**: China

**Contact Person(s)**:

Wang Wenduo, 18907178900@189.cn

Jiao Hailong,18907100678@189.cn

Zhou Huiping, hbzhouhp@chinatelecom.cn

## 1 Use Case Summary Table

| Item | Details |
|---|---|
| Category | 5G |
| Problem Addressed | Traditional natural language understanding (NLP) technology is difficult to handle increasingly complex customer service business scenarios. In practical applications, a significant amount of human intervention is required as a fallback, leading to issues such as inconsistent service quality, repetitive tasks, low service efficiency, and high operational costs for enterprises. |
| Key Aspects of Solution | Introducing the AGENT architecture to build the "Diting" intelligent customer service agent: Using Large Language Model (LLM) for intent recognition, entity recognition, sentiment analysis, dialogue analysis, business opportunity mining, and more. Leveraging the knowledge base to store customer service marketing activity data, telecommunication business knowledge data, and other relevant information for product recommendations. |
| Technology Keywords | Large and small model collaboration, Intelligent Customer Service |
| Data Availability | Not Available |
| Metadata (Type of Data) | Text data, audio data |
| Model Training and Fine-Tuning | Fully parameterize the training of large and small models for intent recognition, entity recognition, and incorporate decision-making judgments to enhance the collaborative capabilities of large and small models. The foundational small model used for this task is Tiny-BERT [4] and the large model used for intent recognition is the Telechat32B [5] or Qwen8B [6]. These are fine-tuned on training data that consists of dialogues between users and customer service agents.Inference-based large models are used for the automatic optimization and generation of service solutions. |

(continued)

| Item | Details |
|---|---|
| Testbeds or Pilot Deployments | The use case is deployed in the Hubei Customer Service Center of China Telecom Corporation Limited. The Hubei Telecom 10000-system currently supports over 200 business scenarios and nearly 800 business interfaces. |
| Code repositories | Not Available |

# 2 Use Case Description

## 2.1 Description

This case involves the development of the "Diting" Intelligent Customer Service Agent, which uses AI to empower China Telecom's 10000 customer service hotline. It provides users with warm, expert-level customer service while reducing operational costs for the enterprise.

The current goal is to leverage "Diting" to quickly learn and adapt to new businesses, accommodate the complexity of operations and the uncertainty of human-computer interactions, and continuously reduce the volume of cases requiring human customer service intervention.

Innovative Technological Approaches Being Implemented at the algorithm level, knowledge base level, and business level are described below:

**Algorithm Level**: The "Diting" Intelligent Customer Service Agent adopts a system architecture that leverages collaboration between large and small models to adapt to different scenarios. It encompasses the entire lifecycle of AI models, including model training, fine-tuning, inference acceleration, model evaluation, and data feedback loops. In terms of model training, it requires obtaining dialog text between users and customer service representatives, which is then manually labeled by an annotation team. To date, over 300,000 pieces of data have been labeled. During application, the collaboration between large and small models enables rapid recognition of user intent and automatic extraction of key information such as ID numbers, phone numbers, and addresses provided by users. For intent recognition, the small model uses the Tiny-BERT architecture and applies the standard BERT embedding techniques such as Token Embedding, Segment Embeddings, and Position Embeddings **[3]**. The training data consists of approximately 300,000 dialogues between users and customer service agents. The large model used for intent recognition is the Telechat32B or Qwen8B, which shares the same dataset as the small model. It is fine-tuned using full-parameter fine-tuning, trained with the LLaMA Factory tool, and deployed using the VLLM framework to support inference. Both models are hosted on servers and made available for use via API calls. Human customer service agents can communicate with users based on recommended scripts provided by the "Diting" agent, resolving complex customer issues.

**Knowledge Base Level**: Telecom services and the latest promotional activities are input into a knowledge base, which is accessed via API calls for scenarios such as Q&A and product recommendations. This knowledge base architecture is based on a graph-structured text indexing framework. Using large language models and natural language processing technologies, entities and relationships are identified from text data. Extracted entities are used as graph nodes, and relationships between entities serve as connecting lines between

nodes, forming a knowledge graph. This graph structure clearly displays the entities and their interrelationships within documents, providing robust support for subsequent knowledge retrieval and answer generation.

**Business Level**: The Hubei Telecom 10000-system currently supports over 200 business scenarios and nearly 800 business interfaces. The "Diting" Intelligent Customer Service Agent introduces a two-layer knowledge graph to encapsulate the expertise of service experts, precisely matching different intents and providing personalized service solutions for customers. The first layer of the graph focuses on service plans, ensuring accurate execution of service actions during the service process to meet diverse customer needs. The second layer focuses on service components, defining the specific business interfaces required for each service action. All necessary interfaces for current customer service scenarios are registered in the "Diting" agent to ensure on-demand invocation and efficient service execution.

LightRAG is used, which consists of two main stages, index building and query response. Index building involves splitting the original documents into small paragraphs. Using a language model, it extracts key entities and their relationships. Duplicate content is removed to create a more compact knowledge structure. All concepts and relationships are stored as a knowledge graph in a graph database. Each concept and text is also converted into a vector using a vector database lookup for fast similarity searches using mathematical methods. When new data is added, LightRAG uses an incremental update approach to update only the affected parts, avoiding full rebuilds.

The Query Response involves the system analyzing the user's query and identifying two types of keywords, Local Keywords and Global Keywords. This two-step retrieval uses local keywords to find the most relevant content in the vector database and use global keywords to explore related information in the knowledge graph via relationships. The retrieved results are combined into a context prompt and fed into a large language model (LLM). The system organizes the output using a predefined response template, ensuring the final answer is coherent, logical, and complete.

**Partners**

- China Telecom Artificial Intelligence Technology Co., Ltd., provides computing power platforms and some AI capabilities. **[7]**

## 2.2 Benefits of the use case

Reduction in Training Time: Traditional customer service relies heavily on repetitive manual labor, resulting in high resource consumption and low efficiency. In contrast, "Diting" automates the handling of user requests through AI, reducing the workload of customer service personnel. The training period for newly hired agents has been shortened from three months to one week, significantly reducing operational costs.

Improvement in Service Efficiency: This case builds an intelligent customer service system, introducing AI capabilities to assist customer service personnel in handling calls in traditional customer service. By automatically recognizing intents, extracting information, invoking component interfaces (e.g., checking phone bill interfaces), and executing service solutions automatically, it significantly improves the efficiency and quality of customer service. The application of "Diting" has reduced the time spent on information sorting and confirmation by 20%, and the average handling time per operation has been cut by 10%. The time required

for scenario optimization, which previously took three days for one developer and 1-2 days for one tester, has been reduced to just one person completing it within a single day, accelerating product iteration and optimization.

Increase in User Satisfaction: Based on dimensions such as survey questionnaires, online feedback, and usage trends, highly satisfactory results have been achieved. The satisfaction rate from survey questionnaires is 95%, and the online evaluation satisfaction rate from frontline users is 99.1%.

Opportunity Mining: Through the "Diting" Intelligent Customer Service Agent, potential business opportunities are identified online by analyzing customer profiles and combining them with marketing strategies to make personalized recommendations. Real-time data analysis and service solution optimization accurately match user needs with business processes, avoiding resource misallocation. On average, 15,000 potential business opportunities are identified daily, with approximately 10% being valid opportunities.

## 2.3 Future Work

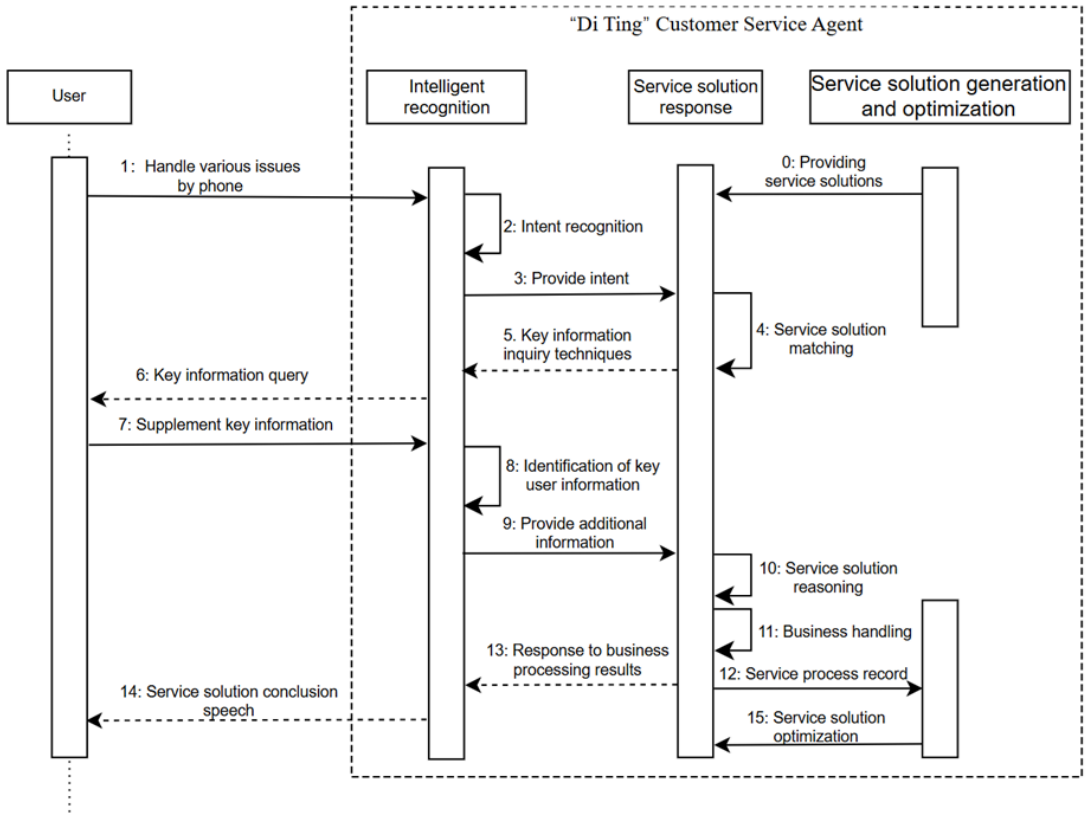
- Further explore the automatic planning and intelligent execution of service solutions: Existing service solutions rely on the experience of business personnel to complete and generate. In the later stage, consider using thinking chains and reinforcement learning to automatically complete or generate service solutions for corresponding scenarios based on real call data to improve the application of AI.
- Expand cross-industry applications: The "Diting" customer service agent has demonstrated strong cross-industry adaptability, and will continue to promote its application in e-commerce, finance and other industries. Customize intelligent customer service solutions that meet the characteristics of the industry with various companies to help companies improve service quality and customer satisfaction.
- Ecological construction and cooperation: The "Diting" customer service agent will actively build a cross-industry ecological cooperation system. By establishing a developer community, opening the API interface and development documents of "Diting", encouraging R&D personnel to carry out secondary development, and jointly promoting the advancement of intelligent customer service technology and ecological prosperity.
- In the future, the "Diting" customer service agent will continue to expand its application scope and market influence, providing intelligent customer service solutions for more industries and helping companies achieve digital transformation and upgrading.
- Transition to Human-Machine Conversation Mode: Further upgrade "Diting" into a voice-enabled intelligent digital employee that directly provides human-to-machine conversation services. The AI can directly engage in conversations with customers, offering multi-round AI interaction capabilities, enhancing contextual understanding, improving the model's ability to soothe customers through conversation, recognizing requests for repetition, and optimizing the interactive experience for customers using online voice self-service. As of June 3, 2025, the voice-enabled intelligent digital employee has been deployed in 18 high-frequency human-machine conversation scenarios, with the rate of complex conversations being transferred to human agents kept below 10%.

# 3 Use Case Requirements

- **REQ-01**: Data Requirements: It is critical to ensure the privacy and security of the data during model training process. The data is sourced from the Hubei Telecom Customer Service Center and consists of voice data between customer service personnel and incoming callers.

- **REQ-02**: Model Capabilities: It is critical to acquire open-source large AI models and conduct secondary training and development to adapt them to specific requirements.
- **REQ-03**: Computing Power Requirements: It is expected that sufficient computing power must be ensured for the inference, fine-tuning, and training of the large model.

## 4 Sequence Diagram




## 5 References

[1] GSMA, Operator Best Practices: AI Large Model Empowering Verticals (Use Cases). [Online]. Available: https://www.gsma.com/about-us/regions/greater-china/wp-content/uploads/2024/12/Operator-Best-Practices-AI-Large-Model-Empowering-Verticals-Use-Cases-CN.pdf

[2] Wolai, "Vertical Industry Large Model Application Scenarios." [Online]. Available: https://www.wolai.com/grk6SvFBakREpFkAKkuTv9

[3] Y. Liu, M. Ott, N. Goyal, et al., "RoBERTa: A Robustly Optimized BERT Pretraining Approach," arXiv preprint, arXiv:1909.10351, 2019. [Online]. Available: https://arxiv.org/abs/1909.10351

[4] Google Research, "BERT: TensorFlow code and pre-trained models for BERT," GitHub repository, 2018. [Online]. Available: https://github.com/google-research/bert

[5] Tele-AI, Telechat: A Telecom Assistant Powered by LLM. GitHub repository. [Online]. Available: https://github.com/Tele-AI/Telechat

[6] QwenLM, Qwen3: Next-generation Open LLM by Alibaba. GitHub repository. [Online]. Available: https://github.com/QwenLM/Qwen3

[7] China Telecom Global, "China Telecom Announces New AI Advancements," [Online]. Available: https://www.chinatelecomglobal.com/sc/whats-new/37325

# Use Case 8: Intelligent Agent Framework and Practice for Network Operations and Maintenance Based on the Collaboration of Large and Small Models

**Organization:** China Telecom Jiangsu Branch

**Country**: China

**Contact Person(s)**:

wenhan Rong, rongwh@chinatelecom.cn, 15301582482

## 1 Use Case Summary Table

| Item | Details |
|---|---|
| Category | 5G |
| Problem Addressed | Network complexity is increasing, with global failures occurring frequently. Approximately 70% of failures are attributed to human factors. The lengthy fault repair time due to multiparagraph troubleshooting and poor planning can result in a significant economic loss. |
| Key Aspects of Solution | • Cloud-Network operation with strategic manual intervention for precise fault resolution.<br>• Deploy digital employees to reduce uncertainties caused by human factors.<br>• Coordinate large and small models to monitor abnormal events, quickly establish an emergency war room, and ensure faults are resolved within 10 minutes. |
| Technology Keywords | Network fault handling, Reinforcement learning (RL), Knowledge Distillation (KD), Agentic AI framework |
| Data Availability | Not Available |
| Metadata (Type of Data) | China Telecom's massive network data and knowledge corpus includes text data, logs, alarms, and configurations. |
| Model Training and Fine-Tuning | Using KD technology to automatically generate Chain of Thought (CoT) corpus in the cloud-network operation field, Supervised Fine-Tuning (SFT), and Group Relative Policy Optimization (GRPO) to train cloud-network operation Large Language Models (LLMs). |
| Testbeds or Pilot Deployments | Not Available |
| Code repositories | Not Available |

# 2 Use Case Description

## 2.1 Description

Global network failures occur frequently, involving multi-paragraph troubleshooting, poor planning, and high processing time for the execution of intelligent agents, which seriously affects user experience. A one-day network outage incurs a loss of 43 billion US dollars.

The solution is composed of the following key components:

(1) Cloud-Network operation Agentic AI: Pioneering a series of LLMs for the cloud-network operation, covering key domains including emergency response, change management, and alarm monitoring. Take the emergency response LLM as an example

- Automated data collection algorithm: Based on historical fault handling reports from emergency incidents, the algorithm extracts data such as fault symptoms (including alarms, logs, phenomenon descriptions, and device configurations), fault automation tools, and fault handling processes. This data is then stored as a knowledge graph to support Retrieval-Augmented Generation (RAG) for agent intent orchestration.
- Adoption of KD technology: After data cleaning and preprocessing, the DeepSeek-R1 (DS-R1) model is employed to complete the reasoning processes underlying fault localization and handling. This augmented data serves as cold-start fine-tuning material, enabling the Qwen model to generate step-by-step reasoning.
- Applying RL-GPRO technology for model fine-tuning to further enhance the accuracy of intent orchestration. The end result is a Cloud-Network Emergency Reasoning Model (CNERM) with a smaller parameter count, capable of generating fault-handling thought processes. For the agent intent orchestration scenario, four reward functions were designed: a step format reward, a step count reward, a step repetition penalty, and an R1 correctness score, ensuring proper model convergence.
- Multi-turn dialogue strategy implementation: Upon detecting a fault, the intelligent agent automatically generates an optimal fault-handling process using the CNERM. Users can refine this plan through multi-turn dialogues. The DS model extracts parameters required for API calls, guiding users to fill in missing information. For critical steps in fault resolution, such as configuration deployment to key devices, where incorrect operations could cause major outages, manual confirmation is required before execution to ensure 100% accuracy.

(2) Standardization of O&M system interfaces: Based on the cloud-network operation low-code platform, establish a universal MCP to connect cloud-network operation data and tools.

(3) Digital employees reduce human uncertainty: Built on the Cloud-Network Agentic AI framework, they monitor over 100,000 network nodes 24/7, replacing manual labor and ensuring consistent, error-free operations.

(4) Large-small model collaboration framework: Rapidly establish an emergency war room to standardize emergency processes. The system automatically creates group chats and invites digital employees, which detect incidents and autonomously invoke emergency response agents. Small models aggregate massive log alarms, while large models perform root cause localization within minutes—achieving the 1-5-10 capability: 1 minute for anomaly detection, 5 minutes for root cause localization, and 10 minutes to resolve faults and restore services. Digital employees handle the entire process automatically and generate intelligent fault-handling reports.

It is critical to realize the whole process control of cloud-network operation emergency events to achieve timely fault discovery, accurate positioning and well-documented disposal. Through

innovation and technology integration, ensure that every emergency can be quickly responded to and properly handled.

Hidden Disease Rectification: A series of automated operations and maintenance tools, such as configuration auditing and log auditing, are generated based on the audit LLM. Incremental tuning of the audit LLM based on real-time network configuration improves the audit accuracy rate by over 95% and further reduces the audit time by approximately 50% compared to the existing baseline.

Timely discovery: Through the synergy of large and small models, real-time monitoring of gold indicators, and rapid identification of abnormal events in cloud-network operation. The next step is expected to build an anomaly dataset covering the entire specialty of cloud-network operation, assess the anomaly detection model more accurately, and improve the model accuracy rate by over 95% through the introduction of time series LLM and model fine-tuning.

Accurate positioning and evidence-based disposal: After emergency events are disposed of, fault cases are automatically stored in the database, and the knowledge graph is automatically generated, so as to realize 100% of the experts' experience in the management and to provide more reference experience for the autonomous learning of the AI agent. The cross-specialty fault location scheme is further refined to ensure that the decision-making accuracy of the AI agent is improved by more than 95%.

## 2.2 Benefits of the use case

The use case focuses on evolving China's Cloud-Network O&M model and has the following impacts:

(1) Reduced network maintenance costs The introduction of digital employees has cut telecom O&M personnel by 30%. Take China Telecom's promotion as an example: 400+ registered intelligent agents, 800+ associated API applications, 800+ daily user visits, and 5,000+ daily usages—saving approximately 5,000+ hours/day on query, emergency, and work order handling.

(2) Enhanced cloud-network operation quality and efficiency: Leveraging existing configuration standards, the system automatically generates configuration templates. In practical production, the abnormal configuration recall rate reaches 94%, precision rate 85%, and China's network device configuration compliance rate has improved to over 95%, ensuring devices "go online without defects" and "cut over without issues".

(3) Improved user satisfaction: 90% of network issues achieve the "1-5-10" fault capability, pioneering nearly one year of "0 public opinion" and "0 failures" for general and above-level faults in the telecom industry. The large-small model collaborative anomaly detection method has reduced original 10-minute detection to 1 minute, with an average accuracy of over 90%. Through intelligent agents for fault localization and handling, fault localization time is shortened by 95%, ensuring fault resolution within 10 minutes and boosting user friendliness to 99.8%.

(4) Cross-industry deployment in China: Deployed in O&M for China's government, culture-tourism, healthcare, and other industries. It has provided innovative digital workforce O&M services to hundreds of clients and trained over 2,000 development-oriented O&M teams.

(5) Implementation in foreign telecom operators: Primarily applied to fault detection and root cause location scenarios.

Awards:

1) " A New Mode and Practice of Human-Robot Interaction Operation and Maintenance Based on Network Large Language Model AI Chat Operations, AIChatOps " selected as a featured case in the ITU AI for Good - Innovate for Impact Final Report 2024.
2) "Intelligent Configuration Auditing" was selected as an industry standard in the Blue Book of Telecom Industry Development.
3) "Artificial Intelligence for IP Network Configuration Auditing" was selected as one of the top ten pioneering cases of "Artificial Intelligence + Telecom Industry" organized by CAICT.
4) "Innovative Applications of China Telecom's Network LLMs in Network Maintenance" certified as an Application Innovation Case by the China Communications Standards Association (CCSA).
5) "Cloud-Network Anomaly Response Center - Watcher" selected as a Starlight Forward - looking Case for Network Intelligence Perception by CAICT 2024.
6) First Prize in the 2024 China Telecom Group Hackathon - LLM Track.

## 2.3 Future Work

To further enhance this use case, future endeavours include:

- Data Collection: Collect a more comprehensive dataset to improve the model and enhance the accuracy of anomaly detection and task planning. Optimize the process of automatic high-quality data collection, enabling the cloud-network operation model to self - learn and improve over time.
- Model Development: Improve Large Language Models (LLMs) to better understand and execute complex network maintenance tasks.
- Extension and Variation: Expand the Agentic AI framework to new scenarios, covering industries other than telecom and government affairs.
- Industry standards: Standardize the full - process solutions of data, models, and tools to provide a reference for network maintenance across industries.

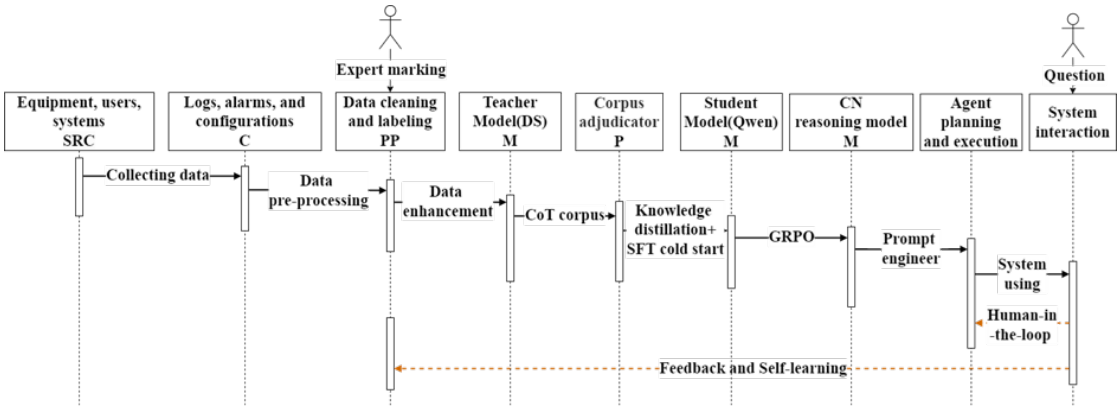
# 3 Use Case Requirements

- **REQ-01**: It is critical to deploy a strategic human-intervention agent framework for deterministic cloud-network O&M.

  Note: The framework allows secondary modification and confirmation of agent intent orchestration and model parameter extraction, avoiding model hallucinations and ensuring 100% execution accuracy.
- **REQ-02**: It is critical to build an automated data collection enhancement algorithm.

  Note: The dataset covers full cloud-network scenarios (such as emergency response, change operations, and alarm monitoring), which can be used to train and fine-tune LLMs, improving model accuracy to over 95%.
- **REQ-03**: It is critical to develop a series of automated O&M tools (e.g., configuration auditing, log auditing) based on large audit models, leveraging real-time network configurations for incremental fine-tuning of audit LLMs.
- **REQ-04**: It is critical to achieve end-to-end control of cloud-network O&M emergency events to ensure timely fault discovery, accurate localization, and well-documented disposal.

## 4 Sequence Diagram




## 5 References

[1] China Academy of Information and Communications Technology, “Leading case of Artificial Intelligence+Telecommunications Industry", CAICT, 2024, [Online]. Available: https://mp.weixin.qq.com/s/7Rndtre-wcromXGCJPPc0w

[2] China Academy of Information and Communications Technology, "Blue Book of Telecommunications Industry Development - Intelligent Development", CAICT, 2024, [Online]. Available: http://www.caict.ac.cn/kxyj/qwfb/bps/202409/P020240926401677616677.pdf.

[3] DeepSeek-AI, "DeepSeek-R1: Incentivizing Reasoning Capabilities in LLMs via Reinforcement Learning", DeepSeek-AI , 2025, [Online]. Available: https://arxiv.org/abs/2501.12948.

[4] Task execution method and device, storage medium, computer program product, R. Wenhan, CN119514687A, 2025. [Online]. Available: https://pss-system.cponline.cnipa.gov.cn/documents/detail?prevPageTit=changgui

[5] Configuration detection method, device, electronic equipment, and storage medium for network equipment, X. Liqin, CN120017511A, 2025. [Online]. Available: https://pss-system.cponline.cnipa.gov.cn/documents/detail?prevPageTit=changgui

[6] Self updating method, device and electronic equipment for log anomaly detection model, Z shu, CN120066896A, 2025. [Online]. Available: https://pss-system.cponline.cnipa.gov.cn/documents/detail?prevPageTit=changgui

# Use Case 9: Smart Assistant for Telecom Business Acceptance Based on the "Hui Xin" Large Language Model

**Organization:** China Telecom Research Institute

**Country**: China

**Contact Person**:

Yingzhao Zhu, zhuyz@chinatelecom.cn

## 1 Use Case Summary Table

| Item | Details |
|---|---|
| Category | 5G |
| Problem Addressed | Due to intricate customer demands, a surge in business volume, cumbersome system operations, high service representative turnover and a large number of new recruits, frontline service representatives at telecom service halls are facing increased work pressure and low service efficiency. |
| Key Aspects of Solution | The smart assistant for telecom business acceptance functions as a personal AI secretary for frontline service representatives at telecom service halls. By processing real-time face-to-face conversations between frontline service representatives and customers, the smart assistant provides comprehensive support functions to help them consistently deliver high-quality customer service, including 360-degree customer profile visualization, customer demand analysis and intent recognition, gold-standard sales pitch generation, instant retrieval and recommendations of telecom service knowledge and AI-powered autonomous orders filling. |
| Technology Keywords | Telecom business acceptance large language model, AI agent technology, natural language processing |
| Data Availability | Not Available |
| Metadata (Type of Data) | Audio speech data, text transcription, and customer-specific parameters |
| Model Training and Fine-Tuning | Speech recognition model, speech enhancement model, speaker separation and recognition model, speech-to-text model, DeepSeek-R1-Distill-Qwen-32B [5], Low-Rank Adaptation (LoRA) |

(continued)

| Item | Details |
|---|---|
| Testbeds or Pilot Deployments | Smart assistant has been deployed on the computers of two China Telecom business halls. It has also been deployed on PC and pads in the laboratory which can assist service representatives with product display, automatic form filling, and SIM card delivery, freeing up 30% of representatives' time for high-value tasks such as customer communication and marketing. |
| Code Repositories | Not Available |

# 2 Use Case Description

## 2.1 Description

For frontline market operations, China Telecom developed the "Hui Xin" LLM (hereinafter referred to as Hui Xin), by leveraging its natural advantages in computing power, data, and technology. Hui Xin is fine-tuned and distilled using massive volumes of telecom business specifications and operational guidelines. It possesses the efficient advantage of Mixture-of-Experts (MOE) **[4]** distributed inference and excels in understanding product marketing, business acceptance, and customer service. This drives the digital and intelligent transformation of China Telecom's various frontline market channels.

Telecom service halls are critical hubs where operators provide a wide range of communication services to customers. However, in recent years, frontline service representatives in these halls face increasing work pressure and declining service efficiency due to complex customer demands, surging transaction volumes, cumbersome system operations, high staff turnover rates, and a large influx of new employees. There is an urgent need to implement artificial intelligence (AI) technologies to assist these representatives in consistently delivering high-quality customer service.

To address this, China Telecom, based on the Hui Xin, has developed a smart assistant for telecom business acceptance **[1]**. This smart assistant breaks through the limitations of traditional Graphical User Interface (GUI)-based acceptance mode, creating a novel Language User Interface (LUI) acceptance paradigm. It provides service representatives with comprehensive AI auxiliary functions, empowering all service representatives to achieve zero onboarding training, gold-standard service quality and hands-free system operations, greatly reducing their work pressure and improving work efficiency.

The Hui Xin-based smart assistant effectively acts as an AI-powered secretary for service representatives. It monitors, captures, and processes the real-time face-to-face dialogue between the representative and the customer. The smart assistant provides the representative with:

- End-to-end processing of real-time conversations: When a customer enters the telecom service hall, the service representative initiates the interaction with a greeting and inquiry about needs. A speech recognition model instantly detects the conversation between the service representative and the customer **[3]**. The speech stream first goes through a speech enhancement model to reduce noise and improve voice quality. Then, a speaker diarization and recognition model separates the speech of the service representative and the customer and identifies which utterances belong to each. Subsequently, a

speech-to-text model transcribes the separated speech into text for each individual, which serves as input to LLM.

- Customer profile visualization and intent recognition: Hui Xin performs semantic understanding and extracts key information, such as customer's phone number. By using this phone number as input field, the AI agent automatically calls the customer information query Application Programming Interface (API) from backend systems and displays the customer profile to the service representative for enhanced customer insight. Concurrently, Hui Xin autonomously identifies the customer's service intent from the real-time conversation. It automatically opens the corresponding system interface and may reduce average customer waiting time by 10%.
- Telecom products recommendation and gold-standard sales pitch generation: After identifying the customer's service intent, the smart assistant dynamically recommends relevant telecom products and generates a gold-standard sales pitch to guide the service representative in customer marketing. The average marketing success rate of service representatives has increased by over 30%.
- Instant retrieval and recommendations of telecom service knowledge: Based on the progress of the conversation between the service representative and the customer, smart assistant also provides real-time prompts about product details or solutions based on dialogue context. Representatives may additionally perform Retrieval-Augmented Generation (RAG) searches for specialized queries **[2]**. This supports service representatives in better answering customer questions, especially important for the novices.
- AI-powered autonomous orders filling: After successful marketing, it may proceed to the system order placement phase and service representatives are no longer constrained by the tedious tasks of the traditional GUI interface, such as extensive manual mouse-clicking to select information and keyboard typing to input customer details. Hui Xin extracts critical information from the conversation and, integrated with Robotic Process Automation (RPA), autonomously populates order forms in the system interface. This GUI form-filling paradigm frees up 20%-30% of sales representatives' time for high-value tasks such as customer communication and marketing, cuts key process error rates by over 50%.

## 2.2 Benefits of the use case

The Smart Assistant enables the creation of an inclusive work environment where all staff at business halls, including novices and individuals with manual dexterity impairments, can receive comprehensive AI assistance. And, the entire process eliminates the need for staff to manually operate cumbersome systems, allowing them to focus on high-value tasks such as customer engagement, marketing, and communication. This ensures that staff can consistently deliver high-quality telecom services without errors. The initiative aims to create more income and employment opportunities for all staff at business halls, including novices and individuals with manual dexterity impairments, thereby advancing decent work and economic growth.

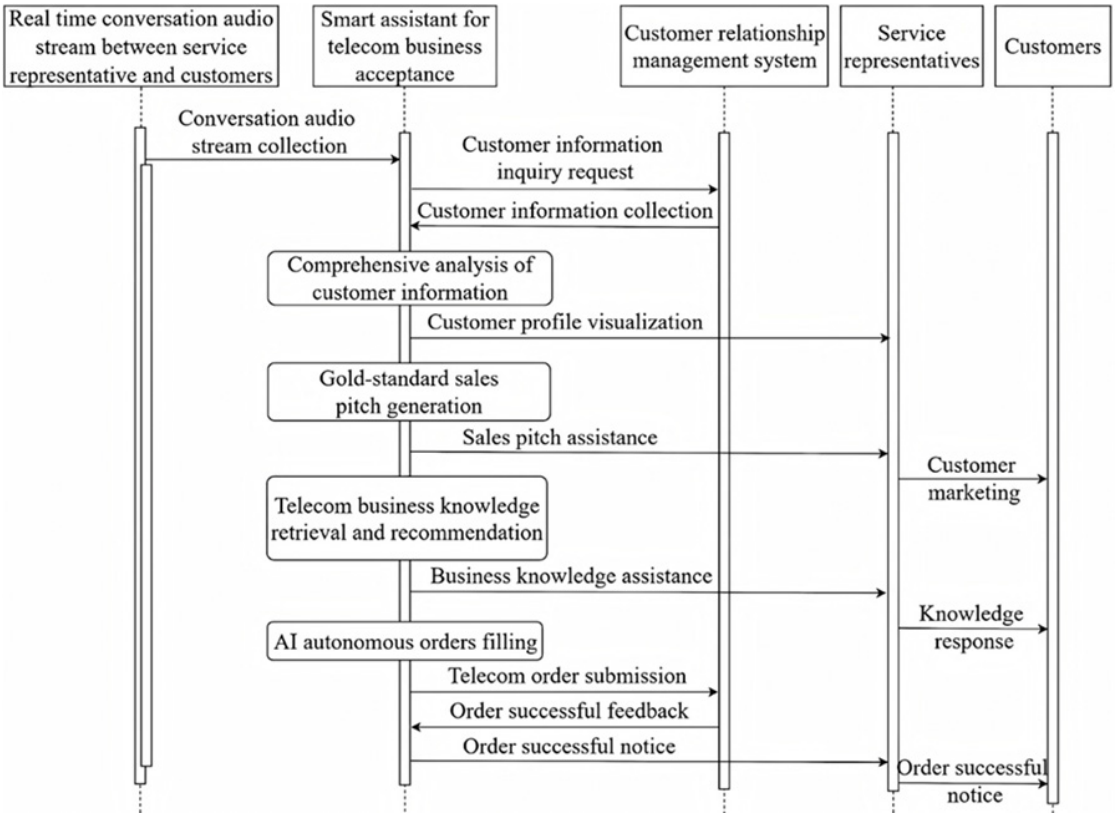
## 2.3 Future Work

In the initial phase, the smart assistant for telecom business acceptance has been deployed on the computers of two China Telecom service halls and demonstrated the ability to support a work environment where all service representatives at service halls, including novices and individuals with manual dexterity impairments, can deliver error-free, high-quality telecom services consistently. Subsequent phases plan to expand the deployment to an additional 100 service halls, with gradual nationwide deployment to all service halls in the future.

## 3 Use Case Requirements

- **REQ-1**: It is critical to fine-tune the telecom business acceptance large language model with internal telecommunication carriers' data, including telecom product and service information and business processes, which enable the large language model to master domain-specific knowledge of telecom service operations.
- **REQ-2**: It is of added value to continuously and iteratively optimise capabilities to adapt to evolving telecom products and business processes.

## 4 Sequence Diagram




## 5 References

[1] H. Zou, Q. Zhao, Y. Tian, L. Bariah, F. Bader, T. Lestable, and M. Debbah, "TelecomGPT: A Framework to Build Telecom-Specific Large Language Models," arXiv preprint arXiv:2407.09424, Jul. 2024. [Online]. Available: https://arxiv.org/abs/2407.09424

[2] G. M. Yilma, J. A. Ayala-Romero, A. Garcia-Saavedra, and X. Costa-Perez, "TelecomRAG: Taming Telecom Standards with Retrieval Augmented Generation and LLMs," arXiv preprint arXiv:2406.07053, Jun. 2024. [Online]. Available: https://arxiv.org/abs/2406.07053

[3] T.-E. Lin, Y. Wu, F. Huang, L. Si, J. Sun, and Y. Li, "Duplex Conversation: Towards Human-like Interaction in Spoken Dialogue Systems," arXiv preprint arXiv:2205.15060, May 2022. [Online]. Available: https://arxiv.org/abs/2205.15060

[4] Mixture of Experts Explained, Hugging Face Blog, Jun. 2023. [Online]. Available: https://huggingface.co/blog/moe

[5] DeepSeek AI, "deepseekai/DeepSeekR1DistillQwen32B," GitHub, 2025. [Online]. Available: https://huggingface.co/deepseek-ai/DeepSeek-R1-Distill-Qwen-32B

# Use Case 10: AI-Driven RF-Based Object Detection and Classification for 5G-Advanced and Beyond

*HFCL logo*

**Organization:** HFCL Limited, Department/Division/Group: 5G BU

**Country**: India

**Contact Person(s)**:

Dr. Sonali, sonali@hfcl.com

Subhas Mondal, subhas.mondal@hfcl.com

## 1 Use Case Summary Table

| Item | Details |
|---|---|
| Category | 5G |
| Problem Addressed | Real-time monitoring using AI-powered RF sensing |
| Key Aspects of Solution | AI enabled Integrated Sensing and communication (ISAC): AI enabled Environmental sensing to create a demography map, AI enabled Object detection and identification |
| Technology Keywords | ISAC, 6G, AI-native RAN |
| Data Availability | Public – mmWave DISC dataset for integrated sensing and communication. Models must be trained across sub-6 GHz and mmWave bands to ensure robustness across propagation characteristics [1]. |
| Metadata (Type of Data) | Numerical. Metadata Descriptions:<br>Each RF sample in the dataset should be tagged with relevant metadata to support model training and validation. This includes:<br>• Environment type: Urban, rural, indoor, outdoor<br>• Scenario: LoS/NLoS, static/dynamic objects<br>• Object properties: Material, shape, motion, size<br>• Signal parameters: Frequency band, SNR, timestamp<br>• Label info: Object class, location, reflectivity<br>Metadata enables supervised learning and helps AI models generalize better across scenarios. |
| Model Training and Fine-Tuning | RF-based object detection is trained using raw reflected signal data across varied environments. CNN-based models are evaluated for classification, with emphasis on differentiating LoS/NLoS scenarios and tracking movement patterns. |
| Testbeds or Pilot Deployments | N/A |

(continued)

| Item | Details |
|---|---|
| Code repositories | N/A |

# 2 Use Case Description

## 2.1 Description

AI-driven RF-based object detection and classification for 5G-Advanced and beyond leverages Integrated Sensing and Communication (ISAC) to enable real-time monitoring and security applications across smart transportation, industrial automation, border surveillance, and next-generation telecommunications. As 5G technology progresses toward 5G-Advanced, integrating object detection and classification into the 5G New Radio (5G-NR) baseband processing framework becomes essential for critical applications. This includes pedestrian safety, where smart transportation systems require real-time object detection to prevent accidents, intruder detection for protecting infrastructure, and border security, where ISAC enhances surveillance and threat detection.

Unlike conventional vision-based detection, which relies on structured image data, RF sensing processes reflected wireless signals that are non-independent and identically distributed (non-IID) due to multipath propagation, interference, and material properties. AI is crucial in transforming RF reflections into structured correlated data for object representation by addressing challenges such as nonlinear distortions, multipath interference, and environmental variations.

The AI processing pipeline begins at the baseband layer, where FPGA-based AI models filter noise, mitigate multipath distortions, and enhance signal clarity. Feature extraction is performed by analyzing delay spread, angle of arrival (AoA), and signal strength variations to create multi-dimensional feature maps for object characterization. CNNs trained on RF feature maps enable real-time inference and mitigation via edge-based tensor processors. Additionally, temporal AI models capture signal variations over time, improving classification accuracy for dynamic objects. AI enables real-time adaptation to diverse conditions, ensuring robust detection in low-visibility scenarios such as fog, darkness, and occlusions where optical sensors fail.

Edge detection in RF-based sensing differs significantly from traditional image processing. While Sobel and Canny filters are commonly used for vision-based systems, RF-based sensing requires alternative methods to define object boundaries. Signal strength variations help estimate object shape, while multi-path reflections provide clues about object edges. Additionally, RF sensing can determine whether an object is hollow or solid by analyzing the reflection pattern—hollow objects produce two distinct edge reflections at the ends, while solid objects generate continuous reflections across their structure. This information enhances the accuracy of AI-driven object classification and environmental mapping.

Implementation and Testing

To validate AI-driven RF sensing, structured RF datasets are collected using anechoic chambers to ensure clean, interference-free data. Additionally, real-world testing is conducted in urban, indoor, outdoor, and high-multipath environments to assess model performance under diverse

conditions. The evaluation metrics focus on detection accuracy, latency, and adaptability, ensuring that AI models generalize well across different RF scenarios.

This solution enhances 6G standardization efforts by optimizing spectrum efficiency, improving network intelligence, and enabling AI-assisted beamforming in massive MIMO implementations. Use cases include V2X communication for autonomous vehicles, intrusion detection for security and defense, and defect detection in Industry 4.0. AI-driven RF sensing reduces hardware costs by leveraging existing telecom infrastructure while improving energy efficiency. Challenges such as computational complexity, the need for extensive RF datasets, and integration with low-latency 5G baseband processing must be addressed.

Conclusion

AI-driven RF sensing frameworks enable real-time, high-accuracy object classification, optimizing 5G-Advanced and 6G ISAC standardization. This research bridges the gap between RF signal processing and deep learning, paving the way for next-generation intelligent communication networks.

AI-based ISAC systems offer real-time intrusion detection, border security, and regulatory compliance.

RF sensing uses delay spread (distance estimation), angle of arrival (AoA for localization), and signal strength (object material and size inference).

Data-Driven AI Training:

To ensure robustness AI models require large, diverse, and high-quality datasets covering:

- Urban and rural settings s.t. model performs well in different propagation conditions.
- LoS and NLoS scenarios to model obstructed and unobstructed RF paths.
- Environmental variations to model robustness to multipath, interference, and fading effects.

An anechoic chamber provides a controlled RF environment where noise and reflections are minimized, to create structured datasets for AI training by:

- Placing RF sensors and transmitters at fixed locations to ensure accurate measurements.
- Introducing various objects of different materials and sizes, such as metal, plastic, and wood, to capture RF signal variations.
- Simulating different environmental conditions, including LoS, NLoS, and controlled multipath scenarios.

The system is sensitive to:

- Material properties: RF reflections vary with conductivity, density (e.g., metal vs. wood vs. plastic).
- Obstacle types: Solid or hollow objects affect signal shape and edge reflections differently.
- Surfaces: Smooth vs. rough surfaces impact signal scattering.
- Interference: Co-channel interference and multipath reflections affect clarity; AI models help mitigate these through adaptive filtering.

AI is integrated into multiple layers of 5G-NR baseband processing to ensure real-time RF-based object detection as

- Preprocessing Layer: The programmable baseband receives communication data and multicore processor is configured w.r.t the data. FPGA-based AI models filter noise, mitigate interference, enhance clarity, and update AI configurations for optimized processing in the next communication cycle.
- Feature Extraction Layer: AI model extracts delay, AoA, and signal strength and converts them into structured feature maps.
- Inference Layer: AI-based CNNs classify objects in real-time using Tensor processor at edge computation.
- AI Edge Compute (Tensor Processor): Processes the extracted features using CNN-based object classification models.

The proposed solution thus performs real-time inference for RF-based object detection and classification.

**Partners**

Potential partner: IIT Delhi

## 2.2 Benefits of the Use Case

The integration of AI-driven RF-based object detection into 5G-Advanced and beyond enhances connectivity, automation, and security across next-generation networks. It also supports smart urban development by improving public safety, mobility, and infrastructure efficiency.

Smart Infrastructure & Cities

AI-powered RF sensing enables intelligent traffic control, vehicle-to-everything (V2X) communication, and real-time surveillance, reducing the need for additional sensor installations. These capabilities improve transportation systems, monitor infrastructure health, and strengthen disaster response. The solution also supports environmental monitoring and promotes optimized resource use in urban environments.

Industry 4.0 & Smart Manufacturing

RF-based AI enables automated quality inspection, predictive maintenance, and robotic operations in low-visibility conditions where optical systems are ineffective. This reduces manufacturing waste and energy consumption, driving more sustainable industrial growth.

Next-Generation Telecom Networks

AI enhances Integrated Sensing and Communication (ISAC), beamforming, and spectrum efficiency, resulting in lower power usage and expanded broadband access. These improvements support digital inclusion and smarter connectivity infrastructure for urban and rural communities alike.

Security & Defense

The solution strengthens urban safety and national security by providing real-time, energy-efficient RF-based surveillance for intrusion detection, border monitoring, and critical infrastructure protection.

Cost-Effective and Scalable Deployment

AI-driven RF sensing reduces infrastructure costs while increasing energy efficiency and scalability. It enables resilient smart city systems that are easier to deploy and manage across diverse urban environments.

## 2.3 Future Work

- Large-scale RF Dataset Development: Building extensive, diverse RF datasets is essential to improve AI model generalization. Future work should focus on creating open, standardized RF datasets covering various environments (urban, rural, indoor, outdoor), object types, materials, and propagation conditions (LoS/NLoS).

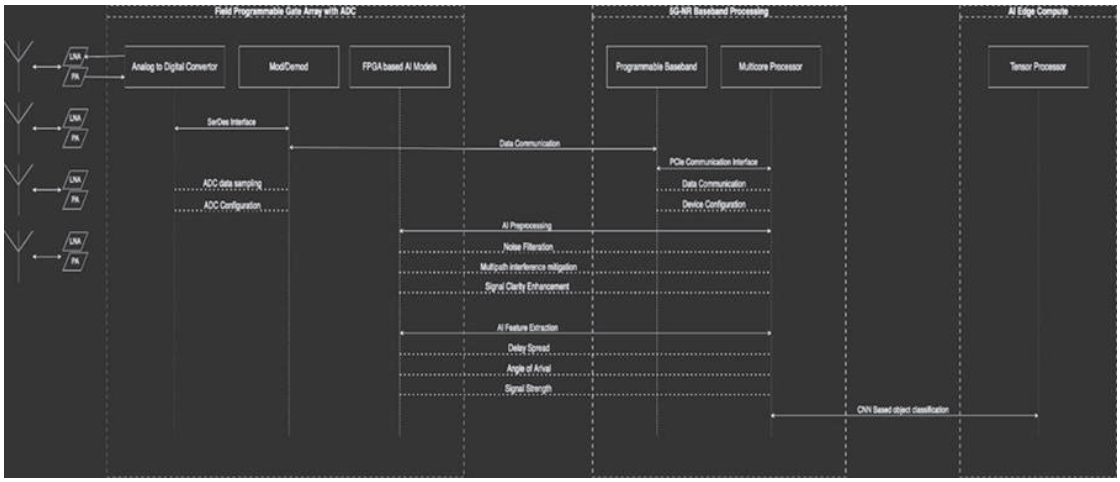
- Optimizing AI Model Complexity: Reducing the computational load of deep learning models without compromising accuracy remains a key challenge. Future research should explore model compression techniques (e.g., pruning, quantization, knowledge distillation) and lightweight architectures optimized for real-time RF sensing.

- Seamless Integration with Low-Latency 5G Processing: Future systems must ensure tight integration of AI-driven object detection with 5G baseband pipelines. This includes designing low-overhead interfaces between AI blocks and physical layer processing, and leveraging AI for real-time beamforming, interference management, and dynamic spectrum access in ISAC scenarios.

# 3 Use Case Requirements

- **REQ-01**: It is expected that the system supports real-time RF-based object detection with latency less than or equal to 1 millisecond.
- **REQ-02**: It is expected that field-programmable gate array (FPGA)-based artificial intelligence (AI) modules are utilized for noise filtering, interference mitigation, and signal clarity enhancement.
- **REQ-03**: It is critical that the system accurately extracts radio frequency (RF) features including time delay, angle of arrival (AoA), and signal strength to enable precise object characterization.
- **REQ-04**: It is expected that edge inference using convolutional neural networks (CNNs) achieves a throughput of at least 100 inferences per second when accelerated with tensor processing hardware.
- **REQ-05**: It is critical that the system maintains robust classification performance in both line-of-sight (LoS) and non-line-of-sight (NLoS) environments, across urban, rural, and low-visibility scenarios.
- **REQ-06**: It is expected that the solution is fully compatible with 5G New Radio (5G-NR) basebands to enable integrated sensing and communication (ISAC) deployment.
- **REQ-07**: It is critical that training and validation are performed using structured RF datasets that include both controlled anechoic environments and real-world deployment scenarios.

## 4 Sequence Diagram




XML file: ML Use Case_AIOD.xml (open in draw.io)

## 5 References

[1] DISC: a dataset for integrated sensing and communication in mmWave systems | IEEE DataPort

[2] Y. Cui, F. Liu, X. Jing and J. Mu, "Integrating Sensing and Communications for Ubiquitous IoT: Applications, Trends, and Challenges," *IEEE Network*, vol. 35, no. 5, pp. 158-167, September/October 2021.

# Use Case 11: AI-Assisted Smart and Efficient Radio Access Networks

**Country**: India

**Organisation**: Indian Institute of Technology Gandhinagar

**Contact Person(s)**:

Ayushman Singh, IIT Gandhinagar, 24310018@iitgn.ac.in

Aman Gupta, IIT Gandhinagar, aman.gupta@iitgn.ac.in

Dr. Sameer G Kulkarni, IIT Gandhinagar, sameergk@iitgn.ac.in

## 1 Use Case Summary Table

| Item | Details |
|---|---|
| Category | 5G |
| Problem to be Addressed | High energy consumption in 5G networks due to redundant paths and over-provisioned link capacities, especially during periods of low traffic demand. Inefficient resource allocation due to a lack of network-wide energy efficiency planning. |
| Key Solution | A dynamic max-min algorithm-based framework on energy-efficient Software-Defined Networking (SDN) that combines network connectivity and maximum network flow with minimum energy consumption, resulting in dynamic end-to-end traffic demands. |
| Technology Keywords | 5G, SDN, Max-Min Optimization, Traffic Prediction, Heterogeneous Cloud Radio Access Network (H-CRAN), Deep First Path Searching Algorithm, etc |
| Data Availability | Mostly, we will rely on open-source data sets. Like cell-level traffic, statistics of 4G/5G. Network Traffic Scenario Prediction, Federated Traffic Prediction for 5G and beyond, Intrusion and Vulnerability Detection in Software-Defined Networks (SDN). |
| Metadata type and granularity | Time-series data: end-to-end traffic demands per time slot (e.g., hourly, daily), link utilization, energy consumption. |
| Model Training and Fine-Tuning | The model can be trained using historical traffic data with daily periodicity to predict future end-to-end traffic demands. Fine-tuning shall be done using real-time network data and feedback from the SDN controller. |
| Metrics and KPIs | Reduction in network energy consumption (%), increased network throughput (%), improved link utilization (%), number of active links, cost reduction (%). |

(continued)

| Item | Details |
|---|---|
| Role of Standards | 3GPP standards for 5G network management, ITU-T Recommendations on energy efficiency in telecommunications. |
| Open Source | Yes, we intend to release the core algorithms and SDN controller modules as open-source. |
| Commercial Utilization | The solution can be utilized and deployed commercially in 5G network infrastructure. |

# 2 Use Case Description

## 2.1 Description

5G networks significantly increase energy consumption compared to 4G, contributing to the ICT sector's growing electricity usage, estimated at 7% globally [1]. A single 5G macro base station may consume ~4.3 kW, nearly 4× that of a 4G site (~1.1 kW), highlighting the urgency of efficient network planning [4]. Traditional energy-saving methods, like shutting down underutilized links, often compromise network performance and fail to respond to dynamic traffic patterns, reducing overall efficiency and increasing carbon emissions. To address these challenges, AI-driven, SDN-enabled solutions are emerging as powerful tools. Organizations like ITU advocate for intelligent planning using big data and machine learning to predict traffic and adapt network behavior in real-time. SDN controllers can dynamically reroute traffic and control base station power states to optimize energy use without degrading service [3]. Our proposed solution adopts a dynamic max-min planning framework for energy-efficient 5G backhaul. The framework balances network connectivity and throughput while minimizing energy consumption by dynamically adjusting network topology and forwarding paths based on traffic demands.

Key Components:

- Traffic Prediction: We leverage historical data and machine learning models to forecast traffic across time slots, enabling proactive planning.
- Topology Adjustment: Based on predictions, the framework activates/deactivates network links to reduce energy usage while maintaining end-to-end connectivity.
- Max-Min Optimization: The planning problem is formulated to maximize network flow while minimizing energy, ensuring scalable and efficient resource allocation.
- H-CRAN Architecture: By separating control and data planes, Heterogeneous Cloud RAN (H-CRAN) supports flexible control, ideal for implementing our framework.
- Distributed SDN Control: A multi-tier SDN controller hierarchy enhances scalability, reliability, and localized decision-making.

Technical Implementation:

- Prediction Algorithm: We will implement a time series or ML-based model to predict end-to-end traffic demands.
- Optimization Engine: The routing problem is modeled as a multi-objective graph optimization task using Max-Min fairness principles, where the aim is to maximize the

minimum link bandwidth utilized while minimizing total energy cost. This is dynamically recalculated as topology changes based on traffic prediction.

- SDN Controller: A distributed SDN architecture dynamically controls switch and base station states based on optimization outputs.
- Simulation & Evaluation: The system will be simulated (e.g., using Mininet), evaluating energy savings, throughput, and link utilization.

Topology Control: The network topology is dynamically adjusted per time slot based on predicted traffic demands. The SDN controller evaluates current link utilization and forecasts, determining which links to activate or put into sleep mode. This uses a real-time link activation matrix that maps traffic load thresholds to energy-optimal link states. Only the minimum number of active links needed for traffic flow and redundancy are kept operational to conserve energy.

Max-Min Routing: Routing is framed as a multi-objective optimization problem that aims to maximize the minimum available bandwidth (max-min fairness) and minimize the energy consumption of active links. This is solved using Integer Linear Programming (ILP) with constraints on link capacities, per-link energy costs, and flow conservation. The optimization is performed periodically (e.g., hourly), with results sent to the SDN controller to update routing and link states. For large-scale settings, a GNN + DRL heuristic is used to learn efficient routing strategies through simulation and real-time feedback.

Path Selection with GNN + DRL:

Traditional methods like LP or weighted Dijkstra are limited in dynamic, large-scale environments. We propose a hybrid ML approach using Graph Neural Networks (GNNs) to represent the network topology and Deep Reinforcement Learning (DRL) to learn adaptive routing policies. This enables real-time decisions that optimize both traffic handling and energy use. Prior studies show that GNN + Distributed RL outperform classical methods in domains like V2X and 5G backhaul [5][6]. To ensure both high throughput and energy efficiency, we combine traditional ILP-based optimization with learning-based GNN + DRL models, forming a scalable and adaptive multi-objective routing solution.

Hierarchical SDN Controller Design:

A two-tier SDN architecture is used. The Global Controller coordinates inter-regional policies and traffic flow. Regional/Local Controllers manage smaller domains like programming data-plane switches and controlling base stations. This structure scales well for dense 5G deployments, supporting quick local responses with centralized oversight.

**Use Case Status:** Initial Prototype

**Partners** : No external partners are currently involved. All R&D is conducted within the 5G Use Case Lab, IIT Gandhinagar. Future collaborations with industry stakeholders and government bodies (e.g., DoT India, BSNL) may be considered.

## 2.2 Benefits of the use case

This solution directly addresses energy efficiency in 5G networks by optimizing resource allocation and dynamically adjusting network topology based on real-time traffic demands. Given the high electricity consumption within the telecom industry, this approach helps minimize energy wastage and contributes to more sustainable network operations.

The system also promotes innovation in network infrastructure by integrating Software-Defined Networking (SDN) and AI-driven optimization. This enables energy-efficient planning and intelligent decision-making, ultimately supporting the evolution of smarter and more efficient 5G infrastructure.

In addition, by significantly lowering energy consumption, the approach contributes to reducing the carbon footprint of telecom operations. The use of a dynamic max-min planning framework empowers network operators to implement greener practices and actively support efforts to mitigate climate change.

## 2.3 Future Work

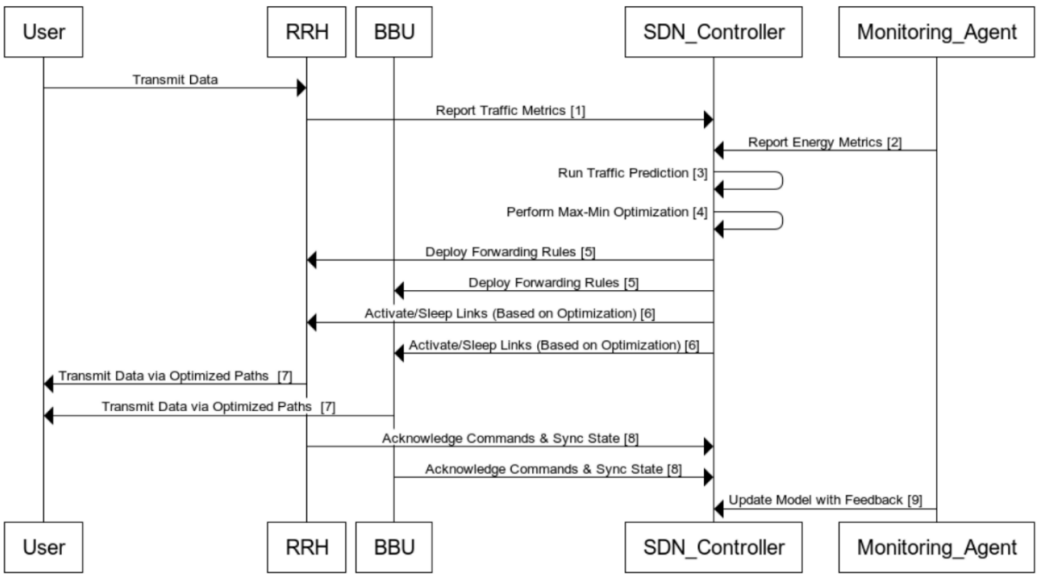
1. Real-Time Implementation and Testing: Transition from framework and idea to simulation and to live testbed deployment, integrating with real/emulated 5G environments.
2. Edge-AI Integration: Explore lightweight AI modules for deployment at edge nodes to reduce latency and increase local autonomy.
3. Multi-Objective Optimization: Extend the current model to optimize for multiple conflicting objectives like latency, throughput, energy, and reliability.
4. Security Integration: Embed security features into the SDN controller for real-time detection of intrusions and malicious traffic in the 5G environment.
5. Scalability to 6G: Investigate how this framework can evolve for 6G technologies, including integration with intelligent surfaces, terahertz links, and satellite backhaul.

# 3 Use Case Requirements

- **UC1**: End-to-end Traffic Monitoring by continuously monitoring end-to-end traffic demands in different time slots to capture dynamic traffic patterns.
- **UC2**: Traffic Prediction: We shall work on employing traffic prediction algorithms to forecast future traffic demands based on historical data.
- **UC3**: Link State Adjustment: Dynamically adjust the active/sleeping state of network links based on predicted traffic demands and network connectivity requirements.
- **UC4**: Max-Min Optimization: Implement a max-min optimization algorithm to determine the optimal network topology and resource allocation for maximizing throughput and minimizing energy consumption.
- **UC5**: H-CRAN Integration: Integrate the solution within an H-CRAN architecture, utilizing the control plane to generate max-min planning and distribute energy-efficient planning schemes to the data plane.
- **UC6**: Distributed SDN Control: Utilize a distributed SDN controller to manage the network, enabling scalability and resilience.
- **UC7**: Time Slot-Based Planning: Implement max-min planning for each time slot based on traffic demand, physical topology, maximum network flow, and minimum energy consumption.

## 4 Sequence Diagram

The sequence diagram illustrates the interaction between network entities, including the user, RRH, BBU, SDN controller, and monitoring agent, highlighting how real-time traffic and energy metrics are used to optimise routing and activate or sleep links for energy-efficient data transmission.




## 5 References

[1] D. Jiang et al., "Rethinking max–min planning on energy-efficient software-defined networking for 5G networks," *Scientific Reports*, vol. 14, 25709, 2024. DOI: 10.1038/s41598-024-76504-9.

[2] United Nations, "Sustainable Development Goals," https://sdgs.un.org/goals

[3] C. Li et al., "Energy-Efficient Traffic Engineering with SDN in 5G Backhaul Networks," *IEEE Communications Magazine*, 2022.

[4] Y. Ding et al., "Carbon emissions and mitigation potentials of 5G base station in China," *Resources, Conservation and Recycling*, vol. 182, 2022, DOI: 10.1016/j.resconrec.2022.106339.

[5] W. Jiang et al., "Graph Neural Networks for Routing Optimization: Challenges and Opportunities," *Sustainability*, vol. 16, 9239, 2024. DOI: 10.3390/su16219239.

[6] Zhang et al., "Distributed Reinforcement Learning with GNNs for Energy-Efficient Network Routing," *arXiv preprint*, 2024. https://arxiv.org/abs/2407.06518v1

# Use Case 12: Dynamic Instance Switching in 5G Network Slices Based on Load

భారతీయ సాంకేతిక విజ్ఞాన సంస్థ హైదరాబాద్
भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

**Country**: India

**Organization**: Indian Institute of Technology, Hyderabad(IITH)

**Contact Person(s)**:

Dr. Antony Franklin (antony.franklin@cse.iith.ac.in)

Mr. Yaswanth Kumar L S (cs24resch11007@iith.ac.in)

Mr. Arjit Gupta (cs23mtech12001@iith.ac.in)

Ms. Somya Jain (cs23mtech12011@iith.ac.in)

Mr. Shreesh Gupta (cs23mtech12009@iith.ac.in)

## 1 Use Case Summary Table

| Category | Telecommunications |
|---|---|
| Problem to be Addressed | Efficient management of 5G network slice instances to handle varying load demands while ensuring optimal resource utilization and Quality of Service (QoS). |
| Key Aspects of the Solution | Implementation of a dynamic management system that monitors network slice loads in real-time and switches between different instances of the same slice based on current demands. |
| Technology Keywords | 5G, network slicing, dynamic instance management, QoS |
| Data Availability | N/A |
| Metadata (Type of Data) | Network traffic statistics, slice performance metrics, resource allocation logs |
| Model Training and Fine-Tuning | Application of ML techniques training on historical network data to predict optimal slice instance configurations under varying load conditions. |
| Testbeds or Pilot Deployments | N/A |

# 2 Use Case Description

## 2.1 Description

In 5G networks, network slicing enables the creation of multiple virtualized networks on shared physical infrastructure, each optimized for specific service requirements such as latency, bandwidth, or reliability. However, fluctuating traffic loads across slices can lead to resource underutilization or congestion, degrading Quality of Service (QoS) for end-users. To address this, dynamic slice instance management systems leverage machine learning (ML) to monitor real-time conditions and autonomously adjust resource allocations [1, 2].

The system employs time series analysis to predict traffic patterns, enabling proactive scaling of slice instances. For example, an ML model might forecast peak usage periods for an enhanced Mobile Broadband (eMBB) slice streaming high-definition video, triggering preemptive resource allocation [1]. A load balancer acts as a proxy, distributing traffic across slice instances and managing their lifecycle—spawning new instances during demand spikes and decommissioning underutilized ones during low load. This ensures each instance operates within its capacity thresholds, preventing overloads.

Load balancer allocates a slice instance to the user if the slice instance can accommodate by analysing the traffic patterns as illustrated in Figure TBD. Load balancer dynamically provisions new slice instances only if the current instances lack the capacity to accommodate incoming UE traffic. This decision is triggered by two scenarios:

1. Real-time overload: Existing instances are nearing or exceeding their throughput limits.
2. Predictive scaling: Machine learning forecasts indicate an imminent traffic surge that would exceed configured capacity thresholds [2].

The load balancer acts as an intelligent proxy, continuously evaluating both live traffic metrics (e.g., packets/second) and predictive models. When thresholds are breached—whether due to actual usage or anticipated demand—it automatically spins up preconfigured slice instances. Conversely, it decommissions underutilized instances during low-traffic periods, ensuring resource efficiency. This dual-condition scaling ensures uninterrupted QoS while preventing premature resource allocation and also aligns with global best practices for autonomous network management, as envisioned in ITU-T frameworks for integrating ML in future networks[1].

Use Case Status: The implementation plan is in progress.

## 2.2 Benefits of the use case

This use case enhances the efficiency and resilience of telecommunication infrastructure by introducing innovative network management strategies. It promotes adaptive and intelligent network behavior through dynamic slice instance management, supporting the continuous evolution of telecom systems.

By fostering innovation in network operations, the solution contributes to the creation of more efficient and resilient infrastructure that can respond to growing demand and operational complexity.

Additionally, the use case supports the development of smart city applications by ensuring reliable and efficient network services. This reliability is essential for enabling real-time urban services, improving the sustainability and livability of modern cities.

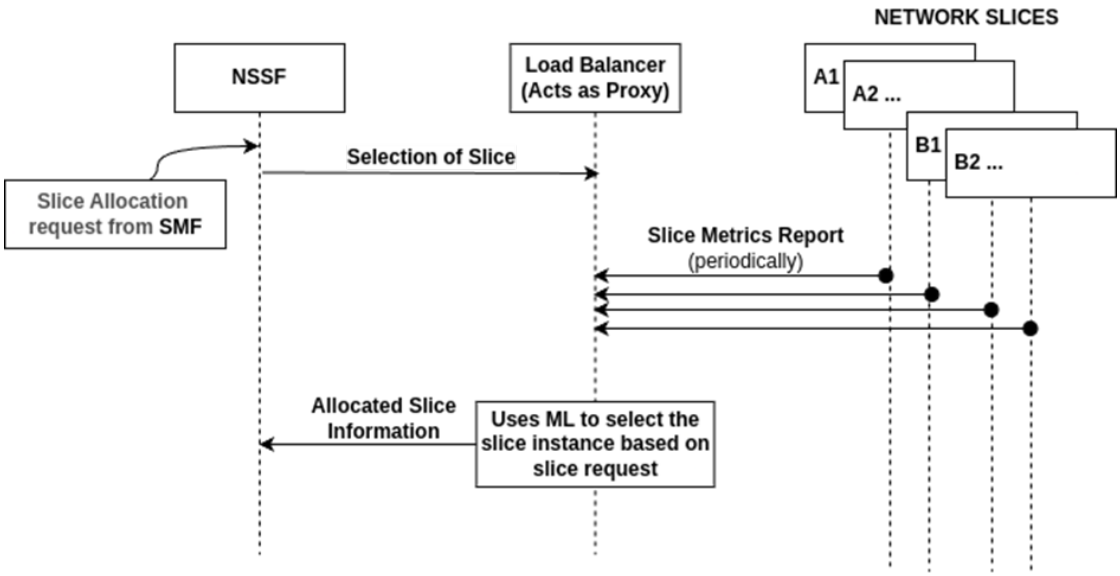
## 2.3 Future Work

- Data Collection: Gather extensive network traffic and performance data to enhance the accuracy of machine learning models (Synthetic data will be used for the ML training).
- Model Development: Develop advanced machine learning algorithms capable of real-time decision-making for dynamic slice instance management by time series analysis of the traffic load.
- Infrastructure Enhancement: Upgrade network monitoring and management systems to support dynamic slice instance switching by the load balancer entity.

# 3 Use Case Requirements

- **REQ-01**: It is critical that the system dynamically monitors load conditions on network slices and adjusts instance configurations in real-time to maintain optimal performance.
- **REQ-02**: It is critical that the solution employs machine learning techniques to accurately predict traffic patterns and proactively adjust slice instances accordingly.
- **REQ-03**: It is critical that the platform ensures dynamic instance switching does not degrade the Quality of Service (QoS) experienced by end-users.
- **REQ-04**: It is critical that the system integrates seamlessly with existing network management frameworks to enable coordinated control and comprehensive monitoring.
- **REQ-05**: It is critical that the solution be inherently scalable to support a large number of network slices with diverse and evolving service requirements.

# 4 Sequence Diagram



## 5 References

[1] Zhao, G., Wen, M., Hao, J., & Hai, T. (2021). Application of Dynamic Management of 5G Network Slice Resource Based on Reinforcement Learning in Smart Grid. In The 10th International Conference on Computer Engineering and Networks (pp. 1485–1492). Springer.

[2] Sattar, D., & Matrawy, A. (2019). DSAF: Dynamic Slice Allocation Framework for 5G Core Network. arXiv preprint arXiv:1905.03873.

# Use Case 13: Intelligent Symbiosis: AI + 5G-A Redefining Sustainable Guardianship in Human-Wildlife Shared Zones

**Country**: China

**Organisation**: China Mobile Communications Corporation, JiuTian Research Institute

**Contact Person**:

Lingli Deng, denglingli@chinamobile.com

## 1 Use Case Summary Table

| Item | Details |
|---|---|
| Category | 5G |
| Problem Addressed | Addressing wildlife protection challenges in areas where human and animal activities overlap. |
| Key Aspects of Solution | Utilizing 5G-A Integrated Sensing and Communication (ISAC) and AI technologies to develop activity detection solutions, reducing human impact on wildlife. Simultaneously employing AI models to enhance supervision efficiency and generate conservation strategies for protected areas, significantly improving conservation outcomes. |
| Technology Keywords | 5G-A ISAC; Multimodal data fusion; Network O&M AI Model; Wildlife Conservation AI Model. |
| Data Availability | Network O&M System; Meteorological Center; Auxiliary equipment (drones, cameras, etc.). |
| Metadata (Type of Data) | Text; Image; Voice; Video; Structured. |
| Model Training and Fine-Tuning | Network AI models: Independently trained and fine-tuned by China Mobile; Wildlife conservation models: Fine-tuned based on open-source models. |
| Testbeds or Pilot Deployments | Commercially deployed in Xiamen, Fujian Province; Pilot deployments in Nanjing and Dalian. |
| Code repositories | Private/internal. |

## 2 Use Case Description

### 2.1 Description

Wildlife population is a key indicator of ecosystem health. However, in areas where human and wildlife activities overlap, human actions pose significant challenges to animal habitats.

Therefore, in such scenarios, it is essential to monitor not only wildlife but also human activities to minimize the impact of human activities on wildlife habitats.

Traditional monitoring of coastal protected areas primarily relies on manual patrols, which are inefficient, limited in scope, and prone to blind spots. Additionally, marine conservation data are diverse and complex, and the current level of data analysis is insufficient to provide accurate information on species distribution, population, and behavior, hindering scientific decision-making for protected area management.

Supported by the UNDP Global Environment Facility, China Mobile developed an "AI + 5G-A Enabled Marine Animal Conservation" solution for Xiamen's wildlife reserve. This solution addresses three key aspects: enhanced vessel monitoring, wildlife data analysis, and conservation strategy generation. Leveraging AI, autonomous networks, and 5G-A Integrated Sensing and Communication (ISAC), it overcomes maritime network challenges while ensuring seamless communication across air, land, and sea. The system provides 24/7 monitoring of vessel trajectories, significantly improving conservation efficiency.

Since its commercial deployment in 2024, the solution has effectively enhanced the habitat of endangered marine species. Compared to traditional methods, the monitored area expanded from 120 $km^2$ to 330 $km^2$, with wildlife distribution increasing by 60%. Rare species sightings and populations have risen significantly. Over 1 million vessel movements are tracked annually, reducing collision-related animal deaths by 97%. Illegal activity detection and enforcement efficiency improved by 45%. Public engagement in wildlife conservation also increased, with discussions on wildlife topics rising by 69%. The solution was awarded the "Outstanding Catalyst - Tech for Good" at the DTW24 Summit in Copenhagen, gaining global recognition.

The solution is highly scalable. As AI and 5G-A ISAC mature, it can be applied beyond coastal areas to rural and forest regions where human-wildlife interactions occur, providing robust support for global conservation efforts.

**Use Case Status**: The use case is part of a larger product development

**Partners** : China Academy of Information and Communications Technology (CAICT),

Xiamen Chinese White Dolphin Amphioxus Nature Reserve Affairs Center, Huawei, ZTE, Inspur, China Unitechs, ZZNode, Xiamen University of Technology

## 2.2 Benefits of the use case

This project significantly improves the working environment and safety conditions for marine wildlife protectors. By enabling remote monitoring from shore-based systems, it eliminates the need for frequent offshore enforcement trips. Over a million vessels can be supervised annually, reducing coastal law enforcement deployments by 50%. This shift allows protectors to operate in a safer and more efficient environment while supporting decent and dignified working conditions.

The project also showcases an innovative application of advanced technologies such as 5G-Advanced and artificial intelligence (AI) in the domain of coastal marine wildlife protection. This not only enhances supervisory capacity within protected areas but also introduces scalable models and new approaches for intelligent infrastructure development across related sectors.

By implementing a precise perception and early warning system to track ship trajectories, the solution effectively prevents unauthorized intrusions into protected areas. This reduces risks to marine life and encourages more responsible practices in maritime activity and resource usage.

Moreover, the project strengthens the overall protection of marine ecosystems by enabling rapid detection and response to threats impacting coastal wildlife. Through the use of intelligent technologies, it helps minimize the negative effects of human activity on marine biodiversity and contributes to a more sustainable balance between development and conservation.

## 2.3 Future Work

In the future, we'll boost the solution's standardization, expand its scope, and promote wide adoption.

For standardization, we'll focus on two dimensions: technological frameworks and interface specifications. Technologically standardized components will enhance solution reproducibility across scenarios, enabling adaptable protection models for diverse endangered species and habitats. Interface standardization will improve cross-vendor interoperability, fostering a sustainable ecosystem that reduces suppliers' R&D costs for scenario-specific customization while lowering integration barriers and procurement expenses in practical deployments.

For scope expansion, firstly, we'll explore AI and 5G-A tech applications in diverse ecosystems, from marine protected areas to grasslands and forests. Based on standardized products, the solution improves applicability and monitoring accuracy in complex environments through scenario-specific tech customization, offering more precise data support for protected area management.

Secondly, we'll promote successful cases from Xiamen Nature Reserve to national and global wildlife conservation. Through international conferences, academic exchanges, and collaborative projects, we'll share technological achievements and management experiences to help more regions improve conservation efficiency and tackle global ecological challenges together.

Lastly, we'll push for the sustainable development of related technologies. We'll strengthen environmental impact assessments of new technologies to ensure their long-term ecological protection suitability. And we'll align tech upgrades with actual needs to avoid over-engineering, ensuring tech development matches conservation goals.

In short, based on this case's achievements, we'll expand cooperation through these efforts and contribute more to wildlife protection, promoting global ecological conservation's sustainable development.

# 3 Use Case Requirements

1) It is critical that the system provides 24/7 monitoring of vessel trajectories, significantly improving conservation efficiency.
2) It is critical that monitored area is very wide with wildlife distribution increasing, considering Rare species sightings and populations that have risen significantly.
3) It is critical that large number of vessel movements are tracked annually, reducing collision-related animal deaths.

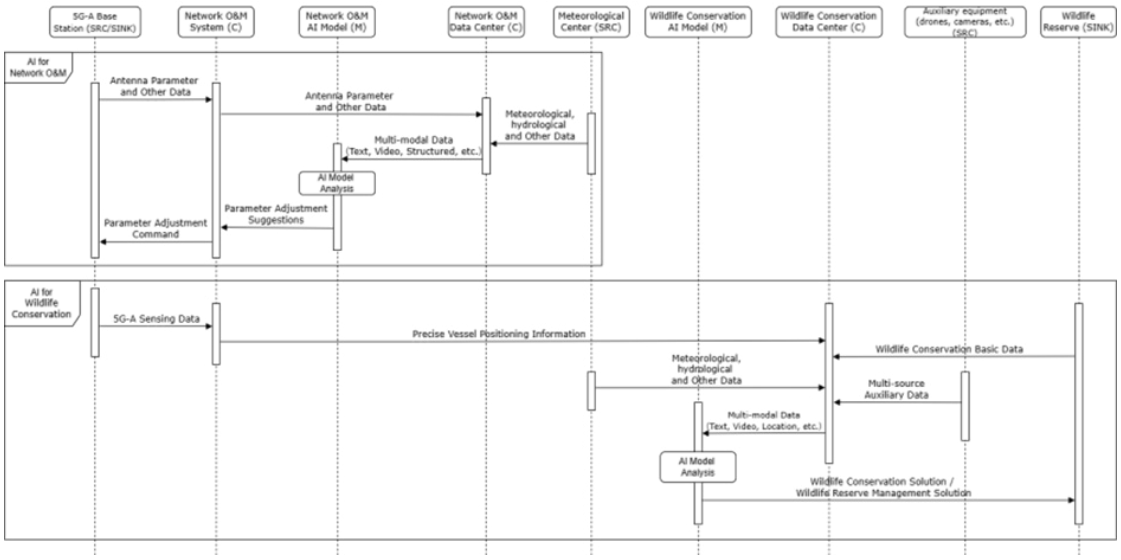
4) It is critical that Illegal activity detection and enforcement efficiency is improved.

5) it is critical that the solution is highly scalable, by applying AI and 5G-A ISAC, applying beyond coastal areas to rural and forest regions where human-wildlife interactions occur, providing robust support for global conservation efforts.

6) it is expected that by enabling remote monitoring from shore-based systems, it eliminates the need for frequent offshore enforcement trips, thereby allows protectors to operate in a safer and more efficient environment.

7) it is expected that by implementing a precise perception and early warning system to track ship trajectories, the solution effectively prevents unauthorized intrusions into protected areas.

## 4 Sequence Diagram




## 5 References

[1] TM Forum, "Wildlife Guardian: AI + 5G Advanced Sustainable Application," Catalysts Project C24.5.747, accessed June 24, 2025, https://www.tmforum.org/catalysts/projects/C24.5.747/wildlife-guardian-ai-5g-advanced-sustainable-application.

# Use Case 14: AI-Powered Network O&M, Ensuring Uninterrupted Network During Power Rationing

**Organization**: MTN Group

**Country**: Southern Africa

**Contact Person(s)**:

Francis Yang, yangmingxi1@huawei.com

Nkosinathi, Nkosinathi.Nzima@mtn.com

## 1 Use Case Summary Table

| Item | Details |
|---|---|
| Category | 5G |
| Problem Addressed | Due to network congestion caused by power rationing, the packet loss rate increases on the bearer network, and the invisible traffic suppression leads to inaccurate network construction, poor user experience |
| Key Aspects of Solution | The main elements of the solution include AI-enabled Network Elements (NE), a network digital map [3][4], and a network optimization agent. An application based on the telecom foundation model [2] enables the automation of network analysis and optimization |
| Technology Keywords | SRv6 (Segment Routing over IPv6), Network Digital Map, NL2SQL (Natural Language to SQL), Network Awareness |
| Data Availability | Not Available |
| Metadata (Type of Data) | Text, including the network object type, performance indicator, and indicator description/ explanation, etc. |
| Model Training and Fine-Tuning | Telecom foundation model based on transformer fine-tuning, 38 billion parameters, natural language corpus |
| Testbeds or Pilot Deployments | Internal Testbed is used inside the organisation |
| Code repositories | Not Available |

# 2 Use Case Description

## 2.1 Description

In the context of the global energy crisis, many parts of the world are struggling with power supply issues, and internet access is frequently disrupted by power outages and restrictions, exacerbating the digital divide in these areas. For example, in South Africa, rotating power outages can sometimes last up to 10 hours **[1]**, causing base stations and routers to shut down and available links to decrease, resulting in network congestion and packet loss that negatively impact user experience. Analyzing individual base stations in manual O&M (Operations and Maintenance) mode requires combining a significant amount of data. However, traditional coding methods are inefficient and inaccurate for data fitting analysis.

AI capabilities of the telecom foundation model and the network data from the digital map are used in the solution to enable the automation of the following network analysis and optimization tasks:

Automatic traffic suppression analysis and real-time visualization:

Based on the network digital map and the network optimization agent's built-in algorithm, the system can collect traffic data at a millisecond level, providing a unique heat map to accurately restore the suppressed traffic distribution of base stations and visually identify network congestion points in real-time. The algorithm models the relationship between packet loss rate and traffic rate, allowing it to calculate the exact amount of suppressed traffic.

Precise global path optimization:

Before any optimization can be effectively applied, it is essential to have network awareness – a real-time understanding of the network topology, link status, bandwidth utilization, latency, and potential congestion points. By continuously sensing and responding to dynamic network changes, the system ensures that optimization decisions are made based on the most current network state. Based on graph theory **[5]** and operations research, a comprehensive algorithmic framework is designed to support a flexible combination of diverse computing policies. And leveraging SRv6 technology, the system enables fine-grained and programmable path control across the IP transport network, allowing services to be steered along optimal paths that meet their specific SLA requirements in various scenarios. This approach not only improves service reliability and performance but also enhances network resource utilization and operational agility.

Predictive operations:

a) Provide minute-level prediction of suppressed traffic through a characterized time-series model for base stations, offering precise capacity expansion suggestions for traffic growth, such as the World Cup.

b) Setting network SLAs in advance, if the packet loss rates or delays of network services reach the warning thresholds due to high link bandwidth utilization or power restrictions, our system will automatically design network paths to enhance service quality. The system offers more than 15 path calculation factors, including bandwidth, latency, and availability, among others. These factors can be flexibly combined to calculate the optimal path. This helps operators classify services more precisely and route them to network paths with different Service Level Agreements (SLAs). As a result, both service-level agreements (SLAs) are met, and resource utilization is maximized.

AI Copilot to simplify network operations:

The integration of an AI Copilot powered by Natural Language to SQL (NL2SQL) technology enables users to perform tasks such as fault diagnoses, configuration changes, and policy recommendations using natural language dialogs. This significantly reduces operational complexity and facilitates the transition from expert-driven to AI-powered network management.

The telecom foundation model is trained on the cloud, and then model-based applications are deployed in MTN. With the AI-based system, the packet loss rate and network delay are significantly reduced, and network utilization is greatly improved. This effectively solves network congestion problems caused by power rationing and line interruption, improves user experience, and further eliminates the digital divide.

**Partners**

- Huawei Technologies **[6]**

## 2.2 Benefits of the use case

The use case deploys AI to enhance quality, reliable and resilient internet infrastructure which can withstand challenges, notably power outage.

Next-generation transport network technologies and agents play a crucial role in boosting network resource efficiency and cutting down device power consumption. AI agents empower intelligent management and optimization of network devices, minimizing the environmental footprint of network operations. This not only advances eco-friendly development in the telecom industry but also strengthens network infrastructure resilience, ensuring uninterrupted and reliable network communication services for all, increase the accessibility to internet in the region and effectively bridging the digital divide.

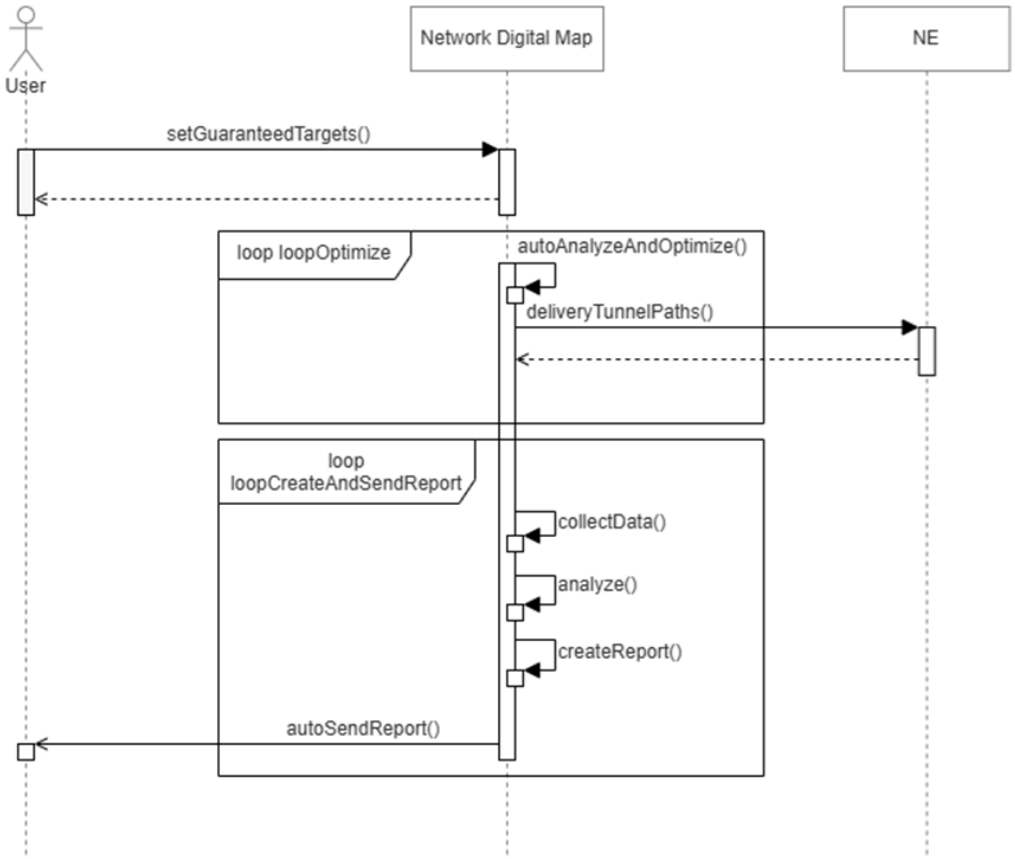
## 2.3 Future Work

Continuously innovate network optimization oriented to the AI WAN. This includes achieving automatic network analysis and optimization based on KPI targets, as well as delivering regular network insight reports that utilize analysis capabilities such as traffic prediction and suppressed traffic visualization. These efforts aim to enhance service experience assurance capabilities and meet the network access requirements of a broader public.

# 3 Use Case Requirements

- REQ-01: It is critical to enable high-precision and high-accuracy SLA measurement for network-wide E2E traffic suppression visualization.
- REQ-02: It is critical to support millions of tunnel management and minute-level network-wide path computation capabilities for large-scale network management.
- REQ-03: It is critical to enable performance instances for network objects to obtain the required performance indicators for network analysis.
- REQ-04: It is critical to provide the appropriate deployment environment for the system, such as servers and computing cards.

## 4 Sequence Diagram




## 5 References

[1] A. Kozak, "Shedding the Load: Power Shortages Widen Divides in South Africa," Commentary (Blog), Kleinman Center for Energy Policy, Univ. of Pennsylvania, 28-Jul-2023. [Online]. Available: https://kleinmanenergy.upenn.edu/commentary/blog/shedding-the-load-power-shortages-widen-divides-in-south-africa/.

[2] "Huawei Builds an AI Model for Telcos – MWC 2024," Aibusiness.com, 2024. [Online]. Available: https://aibusiness.com/verticals/huawei-builds-an-ai-model-for-telcos-mwc-2024.

[3] Huawei Technologies Co., Ltd., IP Network Digital Map White Paper. 2024. [Online]. Available: https://carrier.huawei.com/~/media/cnbgv2/download/products/networks/ip-network-digital-map-white-paper-en-2024.pdf

[4] Staff Reporter, "Empowering Digital Transformation with Network Digital Map," Developing Telecoms, Mar. 27, 2023. [Online]. Available: https://developingtelecoms.com/telecom-technology/telecom-cloud-virtualization/14764-empowering-digital-transformation-with-network-digital-map.html. [Accessed May 31, 2025].

[5] S. Ghosh, A. Mallick, A. Chowdhury, and K. De Sarkar, "Graph Theory Applications in Advanced Geospatial Research," arXiv:2309.03249 [cs.LG], Sep. 2023.

[6] Huawei Technologies Co., "MTN and Huawei signed an MoU to collaborate on the digital future for Africa," Huawei News, Mar. 6, 2025. [Online]. Available: https://www.huawei.com/en/news/2025/3/mwc-mtn-mou-africa. [Accessed: Jun. 8, 2025].

[7] Tmforum, Autonomous networks: Level 4 industry blueprint - high-value scenarios. 2024. [Online]. Available: https://inform.tmforum.org/research-and-analysis/reports/an-level-4-industry-blueprint-high-value-scenarios-11-24

[8] "Draft new Recommendation ITU-T Y.LDT-reqs-funcs: 'IMT-2020 networks and beyond: Requirements and functions for applications demanding large data transmission'," Itu.int, 2020. Available: https://www.itu.int/md/T25-SG13-250725-TD-WP1-0129/en.

# Use Case 15: Digital green transformation of network infrastructure based on artificial intelligence

**Organization**: CHINA TELECOM

**Country**: China

**Contact Person(s)**:

Lika Li, lilk@chinatelecom.cn

Weiye Meng, mengwy@chinatelecom.cn

Miao Yu,yum7@chinatelecom.cn

Jiaming Zhang, zhangjm7@chinatelecom.cn

## 1 Use Case Summary Table

| Item | Details |
|---|---|
| Category | 5G |
| Problem Addressed | The rapid expansion of network infrastructure, such as 5G base stations and data centers, led to a high energy consumption challenge. Infrastructure involving multi-vendor and multi-type devices with inconsistent digitalization and limited energy-saving capabilities necessitates a unified and automated platform integrated by AI and Big Data that can deal with complex problems of frequent faults, inefficient operations, and poor safety assurance in cloud-network operation. |
| Key Aspects of Solution | The use case employs a Cloud-Edge Platform with advanced AI algorithms to handle the complexities. End-to-end integration of multi-type data processing, smart analysis, intelligent control, accurate evaluation, and all-round monitoring enables a comprehensive system solution using ML, DL, RL, and LLM technologies. |
| Technology Keywords | AI energy-saving brain technology, Machine learning, Deep learning, Smart controlling technology and reinforcement learning, Large Model and AI Agent technology |
| Data Availability | Network Operator/Equipment Vendor-sourced Proprietary Data |
| Metadata (Type of Data) | The metadata consists of Wireless network 4G/5G data such as Performance Management (PM), Configuration Management (CM), Call Detail Record, Fault Management, Deep Packet Inspection, Measurement Report, Data Center Monitoring Data (Temperature, Humidity, Load), Cloud servers data (PM, CM, Business), Power data, Operation knowledge. |

(continued)

| Item | Details |
|---|---|
| Model Training and Fine-Tuning | Density-Based Spatial Clustering of Applications with Noise, KMeans, and LightGBM realize profiling analysis of tide, scenario, coverage, and power evaluation.<br>LSTM and CNN predict time-series KPIs (loads, traffic, etc.) using data graphs.<br>Reinforcement Learning techniques such as Q-learning and Butterfly Optimisation Algorithm, optimize threshold and detection upgrades by fine-tuning parameters based on rewards.<br>LLM utilizing CoT, Agent, and RAG frameworks with the DeepSeek model enables smart Q&A, automatic data analysis, and automated operations. |
| Testbeds or Pilot Deployments | The use case has been deployed in 4/5G Base stations and equipment rooms within China Telecom. This use case is part of a large-scale deployment and continuous optimization initiative across China Telecom's national network. |

# 2 Use Case Description

## 2.1 Description

The rapid expansion of 4G/5G base stations, data centers, and equipment rooms has led to soaring energy demands, which are exacerbated by fragmented, multi-vendor device management, low digitalization levels, and unpredictable service loads. Traditional energy-saving methods often struggle with high fault rates, slow emergency response times, and limited operational efficiency, hindering progress toward sustainable digital infrastructure.

To overcome these challenges, a unified AI-powered platform has been deployed to achieve real-time energy optimization and intelligent operations and maintenance (O&M) across complex network environments. The platform integrates a centralized decision-making engine with edge control systems, enabling adaptive, automated control of diverse infrastructure devices, including base stations, servers, and cooling systems.

The key innovations include:

Adaptive Energy Control: The platform continuously aggregates real-time data, enabling dynamic adjustment of device operating thresholds and cooling levels. It automatically identifies high-load scenarios and implements energy-saving measures based on service demands.

Intelligent O&M and Automation: AI-powered agents, supported by modular tools and a domain-specific knowledge base, deliver predictive maintenance, automated inspection, anomaly detection, and real-time reporting **[3]**. This reduces fault resolution time from days to minutes and increases operational efficiency by 90%.

Edge-Cloud Collaborative Control: Edge control engines and cloud decision systems coordinate multi-device operations, optimizing energy consumption while ensuring service continuity and stability. This overcomes the limitations of manual operations and static controls, delivering consistent performance even under variable load conditions.

Scalability and Adaptability: The platform supports integration of heterogeneous devices from multiple vendors. Modular architecture and flexible APIs enable easy adaptation and expansion as service demands evolve, ensuring the system's long-term viability.

Deployed across China Telecom's national network, managing 5.13 million 4G/5G sectors (including 2 million 5G), 3,800 equipment rooms, and 46,000 servers. Achieving over 900 million kWh in annual energy savings and reducing $CO_2$ emissions by more than 500,000 tons. This project demonstrates a scalable, replicable model for achieving large-scale, sustainable digital infrastructure. It highlights the transformative role of AI-powered systems in enhancing operational efficiency, reducing energy consumption, and supporting digital innovation within the ICT sector.

**Partners**

The project involves collaboration with key partners across the ICT ecosystem, including leading equipment providers such as Huawei, ZTE **[1]**, CICT Mobile, and Ericsson **[2]**, which contribute to standardizing and developing energy-saving functionalities, as well as supporting device maintenance and deployment. Integrated IDC service providers, including China Telecom Tianyi IoT and China Telecom Tianyi Cloud, contribute expertise in data center design, construction, and commercial services. Telecom operators such as China Telecom and China Unicom **[8]** oversee large-scale information infrastructure and serve as primary clients for AI-driven energy efficiency initiatives.

## 2.2 Benefits of the use case

1. The project significantly improves energy efficiency and green operations of information infrastructure through the application of AI, big data, and other digital technologies. It ensures the green and high-quality development of next-generation infrastructure while reducing annual greenhouse gas emissions by over 500,000 tons, contributing meaningfully to global climate change mitigation and environmental sustainability.
2. The project has established one of the world's largest cross-disciplinary AI-powered energy-saving platforms, enabling large-scale, safe, and efficient energy optimization across multi-vendor, multi-domain information infrastructure. It reshapes green operation and maintenance processes, reducing deployment and operational costs of energy-saving systems by 90%, and empowering enterprises in their green transformation, cost reduction, efficiency improvement, and sustainable development.
3. Through massive-scale green digital transformation and AI-driven upgrades of information infrastructure, the project drives upstream and downstream industrial upgrades, including for equipment manufacturers and service providers, promoting a green and sustainable transition across the global information and communications technology (ICT) sector.

## 2.3 Future Work

From 2025 to 2027, this project will introduce large language models to reconstruct the AI energy-saving brain, enhance knowledge cognition and intelligent decision-making capabilities, realize self-intelligence and energy savings, and continuously expand the scale of platform applications while improving energy efficiency. Aligned with ITU's mission for sustainable ICT development(e.g., Recommendation ITU-T L.1470 on AI for energy efficiency in networks), it is expected to cover more than 99% of China Telecom's base stations, 95% of class AB IDC business server rooms, and more than 50% of cloud hosts by 2025, saving more than 1.5 billion kWh of electricity per year, which is equivalent to reducing carbon dioxide emissions

by 850,000 tons. Globally, it will disseminate methodologies to establish AI-driven green assessment/optimization frameworks for the full lifecycle of network infrastructure, setting an industry benchmark for sustainable intelligent infrastructure management for broader societal benefits, including climate change mitigation.

Also, this project will extend its impact to empower green transformation across domestic and international industries, enriching AI + low-carbon applications in sectors ranging from smart lighting and green buildings to campus commercial hubs, energy systems, and industrial manufacturing.

## 3 Use Case Requirements

- **REQ-01**: Device Data Collection and Integration

  It is expected that the infrastructure will support standardized data acquisition and ingestion for devices within information infrastructure (such as data centers, base stations, and equipment rooms), covering configuration, performance, energy consumption, temperature **[6]**, and service-level KPIs, both from the device side and the network management (NMS) side **[7]**. The data collection and presentation follow industry standards, ensuring high generalizability and interoperability.

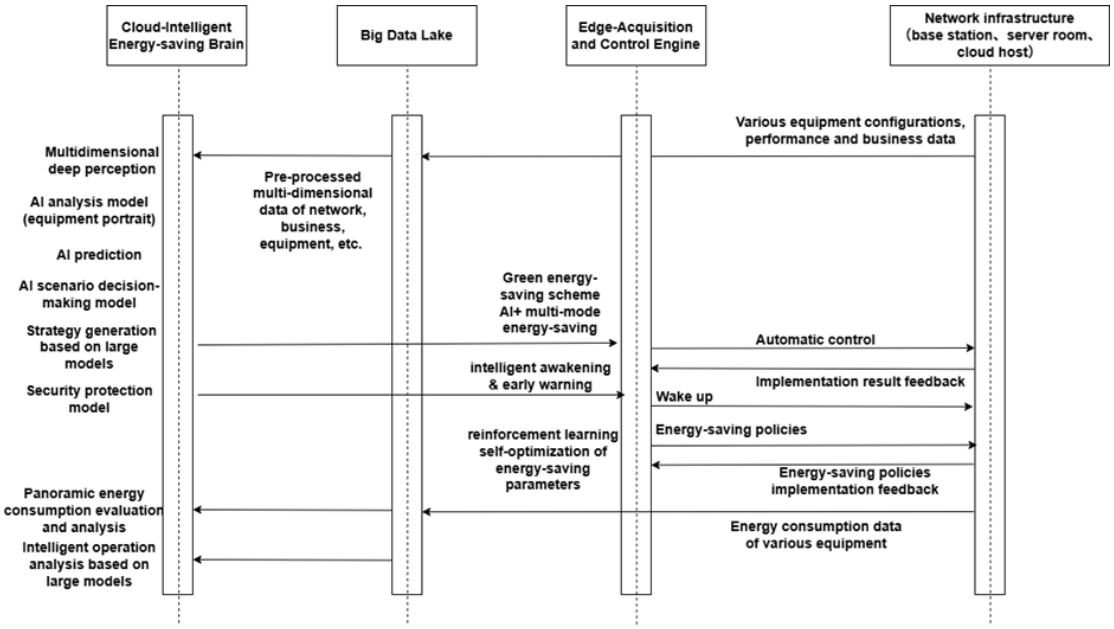
- **REQ-02**: Device/Network Management Control Interface Specification

  It is expected that energy-saving functionalities, instruction sets, parameters, and open interfaces for various devices will be defined. Note: For example, base stations should support symbol shutdown, channel shutdown, deep sleep, and extreme shutdown. Cloud servers should support virtual machine migration, host power-off, and resource downscaling. Air-cooled systems should support switching on and off, temperature tuning, and adjustment of cooling capacity. Water-cooled systems should support startup count control, frequency tuning, and temperature regulation. The platform must support adaptation to standard and non-standard command protocols for heterogeneous devices such as 4G/5G base stations, computing devices (CPU/GPU servers), and cooling systems (air, water, or liquid cooling).

- **REQ-03**: Operational Requirements

  It is expected that during energy-saving operations, the platform must monitor data quality to ensure stable and accurate data streams. It should also continuously track the stability of energy-saving device/environment KPIs, monitor alarm and energy-saving metric changes, and provide mechanisms for manual intervention or emergency wake-up in response to anomalies **[5]**.

## 4 Sequence Diagram




## 5 References

[1] ZTE Corporation, "ZTE – Global Leading Provider of Integrated Information and Communication Technology Solutions," [Online]. Available: https://www.zte.com.cn/global/

[2] "Ericsson – Enabling the full value of connectivity," Ericsson Global. [Online]. Available: https://www.ericsson.com/en

[2] Y. Zhang, X. Wang, H. Zhang, Z. Chen, and Y. Zhu, "Bayesian Network for Base Station Energy Consumption Data Completion," in Proc. IEEE 4th Int. Conf. Electron. Technol. (ICET), Chengdu, China, 2023, pp. 1125-1129. doi: 10.1109/ICET56857.2023.00011.

[3] J. Zhang, Y. Li, C. L. P. Chen, Y. Tang, and H. Liu, "A data-knowledge-driven interval type-2 fuzzy neural network with interpretability and self-adaptive structure," Inf. Sci., vol. 661, Mar. 2024, Art. no. 120173. doi: 10.1016/j.ins.2024.120173.

[4] S. Yang, T. Li, S. Pang, and J. Fujita, "Q-Rung Orthopair Fuzzy Petri Nets for Knowledge Representation and Reasoning," IEEE Trans. Knowl. Data Eng., early access, Jun. 21, 2023. doi: 10.1109/TKDE.2023.3288240

[4] Y. Zhu, Z. Chen, X. Zhang, Y. Wang, X. Wang, and H. Zhang, "An AI Model Automatic Training and Deployment Platform Based on Cloud Edge Architecture for DC Energy-Saving," in Proc. IEEE 6th Int. Conf. Electron. Technol. (ICET), Chengdu, China, 2024, pp. 1334-1338. doi: 10.1109/ICET61505.2024.10579813.

[5] Z. Chen, Y. Zhu, X. Zhang, Y. Wang, X. Wang, and H. Zhang, "Intelligent Data Center Safety Status Prediction based on Algorithm Ensemble," in Proc. IEEE 6th Int. Conf. Electron. Technol. (ICET), Chengdu, China, 2024, pp. 1319-1323. doi: 10.1109/ICET61505.2024.10579810.

[6] Y. Wang, X. Zhang, Z. Chen, Y. Zhu, X. Wang, and H. Zhang, "Exploiting Implicit Neural Representation for Temperature Distribution Reconstruction in Data Centers," in Proc. IEEE 6th Int. Conf. Electron. Technol. (ICET), Chengdu, China, 2024, pp. 1349-1353. doi: 10.1109/ICET61505.2024.10579815.

[7] Z. Chen, X. Zhang, Y. Zhu, Y. Wang, X. Wang, and H. Zhang, "Smart DC: An AI and Digital Twin-based Energy-Saving Solution for Data Centers," in Proc. IEEE 2nd Int. Conf.

Comput. Commun. Artif. Intell. (CCAI), Guiyang, China, 2022, pp. 108-112. doi: 10.1109/CCAI53970.2022.00028.

[8] China Unicom (Hong Kong) Limited, "China Unicom Global," Corporate website, 2025. [Online]. Available: https://www.chinaunicom.com.hk/en/global/home.php

# Use Case 16: AI-Driven Predictive Beamforming for Climate-Resilient 5G/6G Networks

**Organization Name:** Addis Ababa Science and Technology University

**Country**: Ethiopia

**Contact:**

Bereket Teklew Kibret, Bereketteklew@gmail.com

## 1 Use Case Summary Table

| Category | 5G |
|---|---|
| Problem Addressed | Millimeter-wave (mmWave) 5G/6G communication systems are susceptible to significant signal attenuation due to adverse weather conditions, such as heavy rain and fog, resulting in reduced signal attenuation. |
| Key Aspects of Solution | The use of predictive AI algorithms for weather data analysis, dynamic beamforming adjustments, and reinforcement learning-based real-time adaptation enables proactive decision-making and improves signal strength |
| Technology Keywords | 5G, 6G, AI, Predictive Beamforming, Reinforcement Learning, Weather Attenuation |
| Data Availability | The Weather Dataset from Kaggle is used for weather type classification tasks [1] |
| Metadata (Type of Data) | Time-series weather data (rain rate, humidity, wind speed, temperature); numerical beamforming parameters (RSSI, SINR); geospatial data (latitude, longitude, elevation) |
| Model Training and Fine-Tuning | Recurrent Neural Networks (RNN) for weather forecasting, ensemble models for anomaly detection, and Reinforcement Learning (RL) for adaptive beamforming. Training on fused multi-modal weather and signal data, with fine-tuning via transfer learning and policy optimization. |
| Testbeds or Pilot Deployments | The use case is part of a larger research project |
| Code Repositories | Not Available |

# 2 Use Case Description

## 2.1 Description

This use case introduces an AI-driven predictive beamforming system that fuses real-time meteorological feeds with adaptive network control to mitigate weather-induced signal degradation in mmWave 5G and future 6G environments. The system utilizes Recurrent Neural Networks (RNNs), specifically Long Short-Term Memory (LSTM) models, to analyze time-series weather data such as rain rate, humidity, wind speed, and temperature. These RNNs are trained on historical datasets where both weather patterns and corresponding signal quality metrics (e.g., RSSI, SINR) are available. By learning temporal dependencies, the RNNs can forecast short-term weather changes and infer the likely impact on signal attenuation based on empirical relationships. For example, a predicted spike in rain intensity can be associated with a proportional decrease in SINR, allowing the model to estimate future link degradation.

Once the RNN identifies a high probability of signal degradation, ensemble learning methods (e.g., gradient boosting or random forests) are employed to cross-validate the forecast and flag any anomalies such as abrupt changes in channel conditions not accounted for in the weather model. Upon confirming a degradation event, the system transitions to a Reinforcement Learning (RL) phase to determine the best counteractive measures. This ensures that beamforming parameters are not adjusted based on single-model predictions alone but are supported by a multi-model decision-making process.

The RL component is built using Proximal Policy Optimization (PPO) [8], a policy gradient algorithm known for its stability and sample efficiency. The RL environment is a simulated mmWave network where the agent receives observations including current weather state, CSI, and per-beam signal quality. The action space includes beam steering angle, beam width, transmission power, and rerouting decisions. The policy maps environment states to actions that optimize connectivity. The reward function is designed to maximize signal reliability (e.g., maintaining SINR above a defined threshold), minimize power usage, and reduce latency. Negative rewards are assigned for dropped links or excessive energy expenditure. Over time, the RL agent learns to make preemptive, context-aware adjustments that improve network robustness in adverse weather conditions [6].

Actions based on model inference include:

- Beam widening is selected in the action space when rain fade is predicted and tighter beams are likely to fail.
- Frequency shifting is performed to avoid bands more susceptible to atmospheric absorption.
- Power ramp-up, within regulatory limits, is chosen in low-SINR states as a compensatory measure.
- Signal rerouting is triggered when the RL model predicts better connectivity in adjacent cells. In severe cases where no optimal action exists, the system issues operator alerts, representing a fallback policy for manual intervention.

This AI-based approach delivers context-aware, self-healing adjustments that enhance adaptability, reduce downtime, and improve spectral efficiency. It depends on robust, interoperable weather and link data as well as sufficient compute resources, challenges that

will be mitigated through data-validation pipelines and edge inference strategies, providing a scalable foundation for climate-resilient 5G/6G networks aligned with ITU's IMT-2030 vision.

**Partners**

International Telecommunication Union (ITU): Provides the foundational propagation models (Recommendations P.530 and P.840) and guidelines for network resilience.

## 2.2 Benefits of the use case

The proposed system leverages real-time meteorological data, RNN-based forecasting, and RL to proactively mitigate signal degradation caused by weather events. This innovation enhances the resilience and adaptability of telecommunication infrastructure, especially in climate-vulnerable regions. By introducing dynamic and intelligent control mechanisms, the system promotes technological advancement, digital transformation, and robust industry growth.

By maintaining reliable network connectivity during adverse weather conditions, the system safeguards access to critical services such as emergency response systems, telemedicine, online education, and smart transportation. This ensures that both urban and rural populations remain connected and resilient, even during extreme weather events. The ability to maintain uninterrupted communication contributes to urban sustainability and community preparedness.

Weather-induced signal attenuation is a direct impact of climate variability, particularly in the use of high-frequency mmWave bands. By integrating AI models that adapt to and mitigate these effects, the solution acts as a form of climate adaptation technology. It supports the continuity of critical infrastructure under climate stress, helping to reduce system vulnerability.

The proposed system is inherently collaborative, relying on data and coordination between telecom providers, meteorological services, AI researchers, and public agencies. This cross-sectoral integration exemplifies the call for multi-stakeholder partnerships to leverage technology for sustainable development. By aligning telecommunications resilience with environmental data and national digital goals, the project fosters global cooperation, capacity building, and knowledge sharing.

## 2.3 Future Work

Model Development: Building on the initial concept guided by ITU's P.530 [3] and P.840 [2], we will enhance system accuracy by fusing physics-based propagation models with advanced AI techniques. Forecasting accuracy for weather-induced signal degradation will be improved using multi-modal inputs (3D Doppler radar, NOAA weather feeds, IoT sensor data). Metadata extracted from these sources (timestamp, GPS coordinates, elevation, precipitation rate, humidity, temperature, wind speed, and per-beam RSSI) will be formatted in CSV for batch training and JSON for real-time streaming.

As preparatory work, we will evaluate the DeepSense 6G Beam Prediction toolkit [4], analyzing its multi-modal data fusion techniques, benchmarking predictive performance, and extracting best practices for feature engineering to inform our model pre-training strategy and metadata schema design.

Reference Tools and Simulation Environment: An open-source toolkit and simulation platform will be developed for reproducible testing and validation. We will use ns-3 with the mmWave

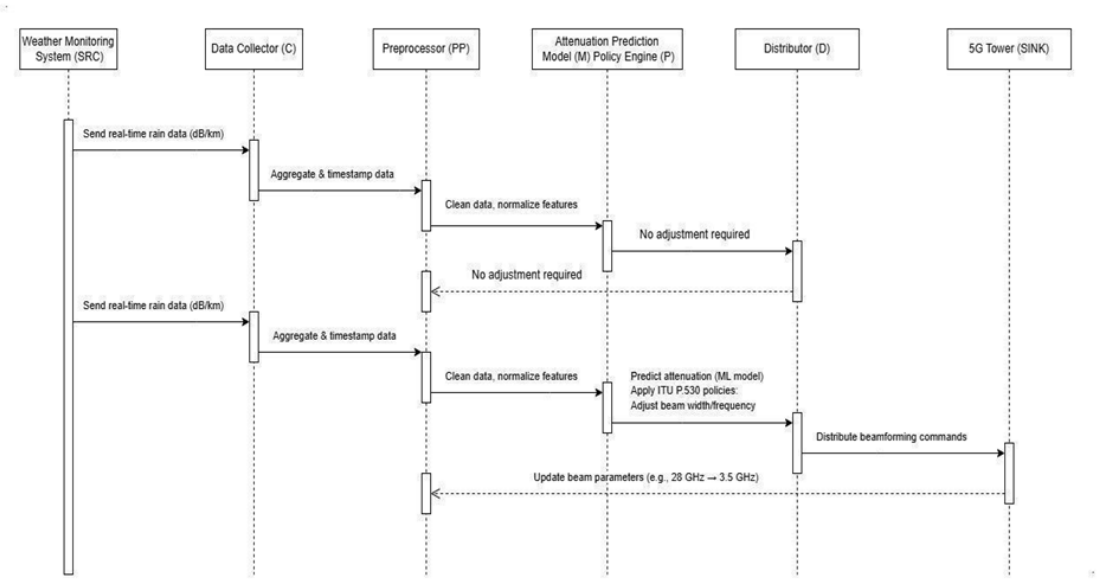
module and WeatherLab toolkit to generate synthetic impairments (rain rates 0–25 mm/h) on virtual cell sites with 64-element arrays [5]. The AI inference pipeline will leverage Python frameworks TensorFlow for RNNs, Stable-Baselines3 for RL policies, and OpenMMLab [7] for data augmentation—to train and deploy models on GPU clusters. All code, simulation configurations, and data schemas will be made publicly available on GitHub under an MIT license.

Standards Development: Close collaboration with ITU and other standardization bodies will be undertaken to establish guidelines for AI-driven network optimization under adverse weather. This will include developing industry benchmarks that ensure seamless integration with existing infrastructures and promote the adoption of resilient communication networks.

## 3 Use Case Requirements

- **REQ-01:** It is critical that real-time weather ingestion is enabled via continuous high-resolution meteorological feeds (e.g., rain rate, humidity, visibility) through APIs or IoT streams.
- **REQ-02:** It is expected that standardized propagation models are incorporated within the inference pipeline to provide baseline estimates of signal impairment.
- **REQ-03:** It is critical that the system captures raw Channel State Information (CSI) including amplitude and phase, along with per-beam RSSI/SINR data from mmWave transceivers.
- **REQ-04:** It is critical that GPU-accelerated edge or cloud compute infrastructure is used to achieve sub-100 ms latency for RNN or reinforcement learning-based inference.
- **REQ-05:** It is of added value that data is structured using standardized JSON or CSV formats with consistent field naming, SI units, and per-epoch manifest records to enhance interoperability.

## 4 Sequence Diagram



## 5 References


[1] N. Patel, "Weather Type Classification," Kaggle dataset, 2023. [Online]. Available: https://www.kaggle.com/datasets/nikhil7280/weather-type-classification/data

[2] ITU-R, Recommendation P.840-8: Attenuation and signal loss in rain for radio communication systems, International Telecommunication Union, Geneva, Switzerland, Sep. 2019. [Online]. Available: https://www.fccdecastro.com.br/pdf/ITU-RP840-8.pdf

[3] ITU-R, Recommendation P.530-18: Propagation data and prediction methods required for the design of terrestrial line-of-sight systems, International Telecommunication Union, Geneva, Switzerland, Sep. 2021. [Online]. Available: https://www.itu.int/rec/R-REC-P.530-18-202109-I/en

[4] DeepSense 6G, Radar-Aided Beam Prediction Toolkit, [Online]. Available: https://www.deepsense6g.net/radar-aided-beam-prediction/

[5] J. Zhang, G. Zheng, Y. Zhang, I. Krikidis, and K.-K. Wong, "Deep Learning Based Predictive Beamforming Design," arXiv preprint arXiv:2302.01035, Feb. 2023. [Online]. Available: https://arxiv.org/abs/2302.01035

[6] F. B. Mismar, B. L. Evans, and A. Alkhateeb, "Deep Reinforcement Learning for 5G Networks: Joint Beamforming, Power Control, and Interference Coordination," arXiv preprint arXiv:1907.00123, Jun. 2019. [Online]. Available: https://arxiv.org/abs/1907.00123

[7] OpenMMLab, "OpenMMLab: Open-Source Computer Vision Algorithm System," GitHub organization, 2018–2025. [Online]. Available: https://github.com/open-mmlab

[8] [1] J. Schulman, F. Wolski, P. Dhariwal, A. Radford, and O. Klimov, "Proximal Policy Optimization Algorithms," *arXiv preprint arXiv:1707.06347*, Jul. 2017. [Online]. Available: https://arxiv.org/abs/1707.06347

# Use Case 17: Anomaly Detection Across 5G Slices Using Federated Learning

భారతీయ సాంకేతిక విజ్ఞాన సంస్థ హైదరాబాద్
भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

**Organization:** Indian Institute of Technology, Hyderabad(IITH)

**Country**: India

**Contact Person(s)**:

Dr. Antony Franklin (antony.franklin@cse.iith.ac.in)

Mr. Yaswanth Kumar L S (cs24resch11007@iith.ac.in)

Mr. Arjit Gupta (cs23mtech12001@iith.ac.in)

Ms. Peddi Manognya (cs24mtech12020@iith.ac.in)

## 1 Use Case Summary Table

| Item | Details |
|---|---|
| Category | 5G |
| Problem to be Addressed | Detecting anomalies in one 5G network slice and preventing their occurrence in other slices by sharing learned model parameters through federated learning. |
| Key Aspects of the Solution | A decentralized anomaly detection system leveraging federated learning to ensure knowledge transfer across slices while maintaining data privacy. |
| Technology Keywords | 5G, anomaly detection, federated learning, network slicing, AI-driven security |
| Data Availability | Self Generated |
| Metadata (Type of Data) | Network logs, slice performance metrics, anomaly detection signals |
| Model Training and Fine-Tuning | Federated learning framework deployed across slices, utilizing anomaly detection models trained on distributed slice data. |
| Testbeds or Pilot Deployments | Open Source tools such as RAN+UE emulator and Free5GC |

# 2 Use Case Description

## 2.1 Description

Modern 5G networks use network slicing to allocate dedicated resources for different services. However, anomalies—such as security breaches, misconfigurations, or unexpected traffic surges—in one slice can affect others. While deploying separate anomaly detection models per slice is possible, their effectiveness is limited. These models rely only on slice-local data, which may miss anomalies that appear similarly across multiple slices.

Data sharing between slices is often restricted due to privacy concerns, and exchanging raw data can lead to redundant training processes. To address these challenges, we propose a federated learning-based anomaly detection framework that enables cross-slice learning without compromising data privacy.

This approach leverages the Network Data Analytics Function (NWDAF) as defined in Section 5 of 3GPP TS 23.288 [3], using standardized interfaces for secure and efficient exchange of model updates. Key features such as message integrity and compression ensure low-overhead and trustworthy communication. When an anomaly is detected in one slice, its signature is generalized through federated learning and shared across all participating slices as illustrated in **Figure TBD**. This enables real-time anomaly detection and mitigation in other slices, even before similar issues occur for which the local model isn't trained to detect.

The framework maintains strict data isolation, reduces detection latency, and strengthens overall network reliability. It aligns with both 3GPP architectural guidelines and ITU-T AI integration frameworks, emphasizing privacy-preserving, distributed intelligence in telecom networks.

**Use Case Status:** Implementation in progress.

**Partners :** No partners

## 2.2 Benefits of the use case

This solution enhances 5G network security and resilience through AI-driven anomaly detection, contributing to the development of reliable and sustainable communication infrastructure. By proactively identifying and mitigating disruptions, it ensures the consistent performance of critical network systems.

The system also supports smart city applications by improving overall network reliability and security. This reduces service disruptions and enables seamless connectivity for Internet of Things (IoT) services, autonomous systems, and digital public infrastructure.

In addition, the federated learning-based approach strengthens cybersecurity and data privacy by preventing fraud and cyber threats. It promotes trust and transparency in digital communications while safeguarding sensitive information in critical communication environments.

## 2.3 Future Work

- **Planned real-world deployment:** The framework can be validated using open-source emulators (e.g., srsRAN+srsUE, Free5GC). This deployment will assess the operational

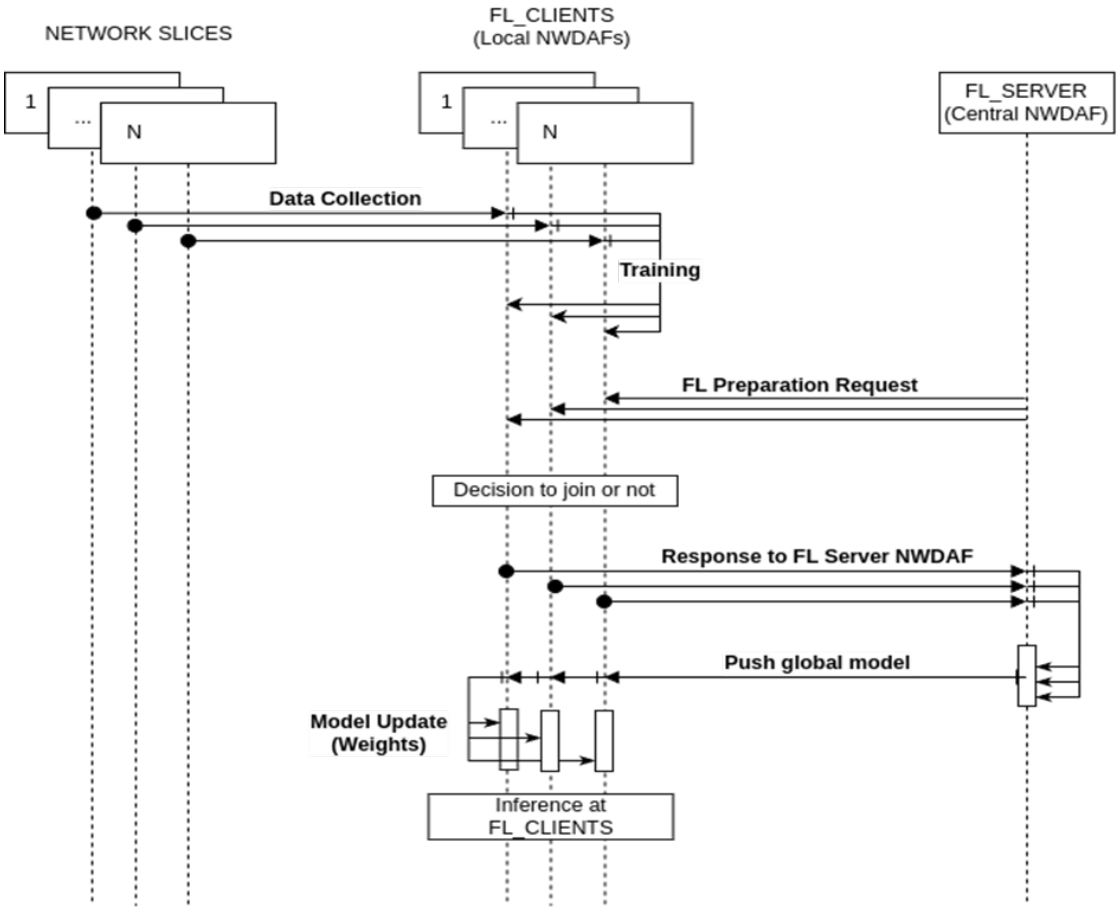
robustness of the federated learning system under realistic traffic loads and network slicing dynamics.

- **Improvements to adversarial and distributed training:** Future work will incorporate robust federated learning methods—such as differential privacy, secure aggregation, and Byzantine-resilient averaging—to defend against poisoning attacks and maintain model integrity in untrusted or distributed environments.
- **Beyond telecom use cases:** The system is extendable to multi-tenant domains like smart grids, industrial IoT, healthcare, and vehicular networks, with initial trials planned in energy and healthcare sectors.

## 3 Use Case Requirements

- **REQ-01:** It is critical that the anomaly detection system operates across multiple 5G slices without directly sharing raw data to preserve data isolation.
- **REQ-02:** It is critical that the federated learning framework effectively generalizes anomaly patterns and distributes model updates efficiently across participating nodes.
- **REQ-03:** It is critical that the system ensures data privacy and adheres to applicable regulatory standards at all times.
- **REQ-04:** It is critical that the solution supports real-time anomaly detection and mitigation while maintaining minimal latency to ensure timely response.
- **REQ-05:** It is critical that the framework is scalable to accommodate dynamic and evolving 5G slice configurations.

## 4 Sequence Diagram



## 5 References

[1] Bonawitz, K., Eichner, H., et al. (2019). "Towards Federated Learning at Scale: System Design."arXiv preprint arXiv:1902.01046.

[2] Zhaohui Yang, Mingzhe Chen, Kai-Kit Wong, H. Vincent Poor, Shuguang Cui, Federated Learning for 6G: Applications, Challenges, and Opportunities, Engineering, Volume 8, 2022, Pages 33-41, ISSN 2095-8099, https://doi.org/10.1016/j.eng.2021.12.002.

[3] Inside TS 23.288: Network Data Analytics Functional Description," Tech-invite.com, 2025. https://www.tech-invite.com/3m23/toc/tinv-3gpp-23-288_c.html.

# Use Case 18: GIS-Integrated Digital Twin for 5G/6G Energy Efficient Base Station

**Organization:** Amrita University

**Country**: India

**Contact Person(s)**:

Prof. Sangeeth Kumar, sangeethk@am.amrita.edu

Prof. Sethu Rao, sethu@am.amrita.edu

## 1 Use Case Summary Table

| Item | Details |
|---|---|
| Category | 5G |
| Problem Addressed | Base stations or cell towers account for 60% to 80% of the total power consumption in a telecom network. Static 5G base station planning leads to energy waste due to signal loss, retransmissions, congestion, and poor performance in dynamic urban environments. |
| Key Aspects of Solution | GIS-based AI Digital Twins of 5G base stations enable low-carbon, energy-efficient 5G networks through real-time optimization. |
| Technology Keywords | GIS-Based 5G/6G Base stations, ArcGIS, AI assisted Digital Twin, Low-Carbon Infrastructure, Genetic Algorithms (GA). |
| Data Availability | OpenRAN<br>Public data from OpenRAN deployment logs and performance traces; additional simulation data from ArcGIS integration. |
| Metadata (Type of Data) | Network Traffic |
| Model Training and Fine-Tuning | Genetic Algorithms (GA), Proximal Policy Optimization (PPO) |
| Testbeds or Pilot Deployments | OpenRAN based 5G Network |
| Code repositories | N/A |

# 2 Use Case Description

## 2.1 Description

The 5G Base stations or cell towers account for 60% to 80% of the total power consumption in a telecom network. This is due to 5G network planning that relies on static configurations and periodic manual updates, which are not suited for rapidly evolving urban environments. Changes in urban infrastructure, buildings, vegetation, and terrain often lead to signal obstructions, signal loss and retransmissions. This causes energy wastage in base stations. These outdated methods not only degrade user experience but also contribute to increased carbon emissions making them unsustainable for future 5G and 6G deployments.

The primary objective is to develop a **self-adaptive, energy-efficient 5G/6G infrastructure** that minimizes carbon footprint while maintaining optimal performance. The use case aims to empower telecom operators, smart city planners, and network engineers with an AI-driven system that can autonomously adapt GIS real-time data to telco base stations for reducing energy wastage.

This solution combines (i) **GIS-based spatial intelligence,** (ii) **AI-powered Digital Twins** that emulates real-time 5G base stations based on GIS intelligence in 5G core for energy efficiency.

(i) GIS-based spatial intelligence:

The 5G Core communicates with ArcGIS using REST-based APIs (using JSON format).

It utilizes the Feature Service (https://<arcgis-server>/arcgis/rest/services/<service-name>/FeatureServer) to fetch data on buildings, terrain, and population density. This data is then used to simulate and optimize base station configurations in real-time. Additionally, the Geometry Service (https://<arcgis-server>/arcgis/rest/services/Geometry/GeometryServer) is employed to perform line-of-sight visibility analysis based on terrain and building data.

(ii) AI Powered Digital Twin

The 5G core emulates connected 5G base stations as digital twins for energy reduction in 5G network. The twin helps in evaluating different base station parameters using ArcGIS, before sending to physical OpenRAN 5G base station. To evaluate different parameters, the twin uses Genetic Algorithms and Proximal Policy Optimization at 5G Core.

1) Genetic Algorithms (GA):

In a 5G core, GA is used to optimize network-wide parameters in base station twin for beamforming and transmission power. ArcGIS has information on terrain, buildings, and population density. The 5G Core using GA and ArcGIS to simulate various network configurations in 5G base station twin. It then evolves to find the most efficient setup.

Optimizing Base Station Placement: GA evaluates the connected base station twin to maximize coverage and minimize interference with energy efficiency as its goal. It ensures that each base station is placed in an optimal position based on the surrounding terrain and building density.

Optimizing Transmission Power and Beamforming: The algorithm uses evolutionary principles to adjust transmission power and beamforming techniques to ensure that signal strength is sufficient in high-demand areas while reducing power usage in areas with low demand.

Once the long-term network configurations are optimized using GA, the twin configurations are sent to the RAN Intelligent Controller (RIC) within the OpenRAN stack. The RIC then communicates these optimal settings to the Distributed Unit (DU) and Radio Unit (RU), ensuring that the network operates efficiently while meeting coverage and capacity requirements.

2) Proximal Policy Optimization (PPO)

While Genetic Algorithms (GA) focus on global optimization and long-term planning, Proximal Policy Optimization (PPO) is employed for real-time dynamic adjustments with the 5G base station. PPO is a type of reinforcement learning (RL) that enables the network to adapt its parameters in response to ongoing Key Performance Indicators (KPIs) such as signal strength and energy consumption. Once the KPIs are maximized, the current configuration is sent as feedback to GA for future changes in GIS. Hence PPO helps in building a closed loop system for digital twin.

The functionalities of PPO are,

Adjusting Transmission Power and Beamforming: In real-time, PPO continuously monitors the network conditions based on setup by GA and adjusts parameters like transmission power. This ensures that signal strength is maintained across areas of high demand while minimizing energy wastage in areas with lower user density or signal requirements.

Efficient Resource Allocation: PPO dynamically manages resource allocation, ensuring that available spectrum and time slots are used efficiently. For example, it reduces transmission power or offloads traffic during periods of low demand, thus optimizing energy consumption.

PPO operates within the RIC in the OpenRAN stack that communicates with the DU and RU in real-time. The PPO adjusts network parameters based on live data and sends these updates immediately to the OpenRAN components for ensuring energy efficiency. At the same time, it sends parameters as a feedback loop to GA algorithm for future optimization considerations.

The proposed solution enables significant reductions in energy consumption and $CO_2$ emissions by avoiding wasteful static configurations and overprovisioning. Moreover, its autonomous, scalable nature positions it as a key enabler for green, sustainable 6G networks, supporting global climate goals and sustainable smart city development.

### Partners

Amrita University, ESRI ArcGIS, Indian Institute of Science, Indian Open Source Mobile Congress

## 2.2 Benefits of the use case

This use case strengthens digital infrastructure by enabling 5G networks to dynamically adapt to environmental changes through the integration of AI technologies and ArcGIS geospatial data. This enhances the responsiveness, efficiency, and sustainability of telecom systems.

By leveraging real-time ArcGIS data, the solution supports the seamless evolution of 5G networks alongside urban growth. It helps reduce connectivity gaps and congestion, directly benefiting smart city initiatives such as Internet of Things (IoT) deployments, autonomous transportation systems, and emergency response networks.

The solution also contributes to environmental sustainability by reducing energy consumption at base stations. AI-driven radio parameter tuning minimizes power usage and lowers the carbon footprint of network operations. ArcGIS integration further supports environmental monitoring for issues such as deforestation and urban heat islands.

Finally, the platform fosters cross-sector collaboration between telecom providers, city planners, AI researchers, and public agencies. By uniting expertise and data from multiple domains, the use case promotes innovation and scalable solutions for smarter, greener, and more connected cities.

### Partners

Amrita University, ESRI ArcGIS, Indian Institute of Science, Indian Open Source Mobile Congress

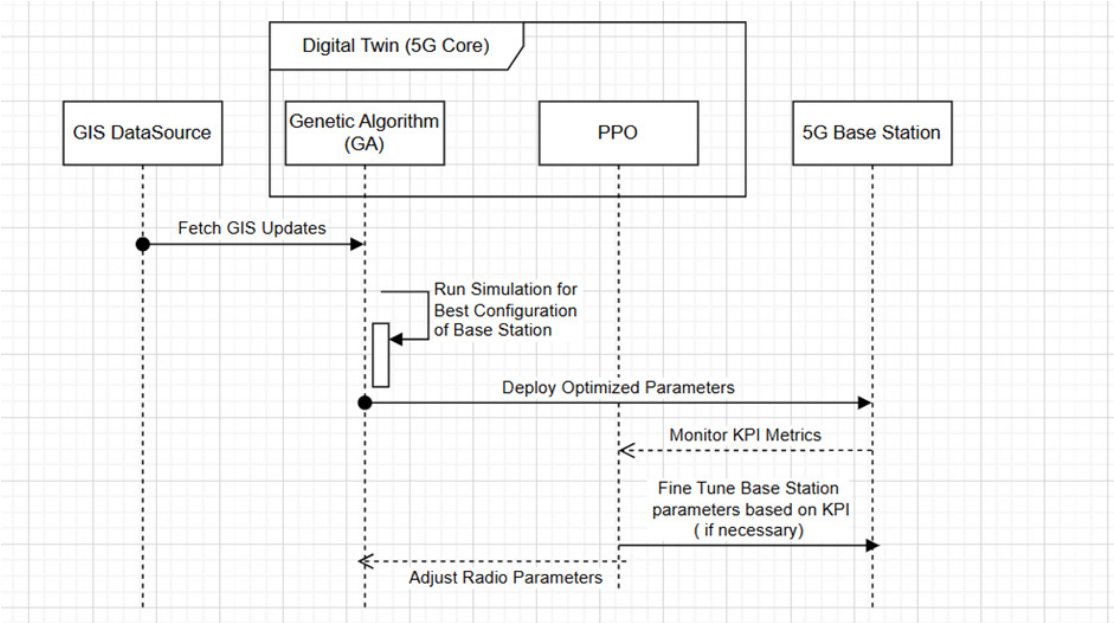
## 2.3 Future Work

- Integrating ArcGIS into 5G core and OpenRAN
- Incorporating Digital Twin for Proof of Concept

# 3 Use Case Requirements

- **REQ-01**: It is expected that the 5G Core system supports real-time geospatial data ingestion from ArcGIS using REST APIs for both feature and geometry services.
- **REQ-02**: It is critical that a continuously updated Digital Twin of each 5G/6G OpenRAN base station is maintained within the 5G Core to simulate network behavior under varying environmental and traffic conditions.
- **REQ-03**: It is expected that the system utilizes Genetic Algorithms (GA) to optimize radio configurations and power allocation in order to minimize energy consumption while maintaining service coverage and signal quality.
- **REQ-04**: It is critical that real-time radio parameter tuning is achieved through Proximal Policy Optimization (PPO), based on live Key Performance Indicator (KPI) feedback from physical base stations.
- **REQ-05**: It is expected that the system operates autonomously with minimal manual intervention, enabling self-optimization and self-healing capabilities across the network.
- **REQ-06**: It is critical that the solution remains fully compatible with Open Radio Access Network (OpenRAN) RAN Intelligent Controller (RIC) standards to ensure vendor interoperability and future-readiness for 6G deployment.

## 4 Sequence Diagram




## 5 References

[1] List academic papers, standards, official reports, technical documentation, or other resources referenced or relevant to the use case.

# Use Case 19: Bandwidth efficient live interaction with virtual 3D demonstrator using semantic communication and GenAI

**Organization:** Tata Consultancy Services

**Country**: India

**Contact Person(s)**:

Abhijan Bhattacharyya, Abhijan.bhattacharyya@tcs.com

Ashis Sau, ashis.sau@tcs.com

Suraj Mahato, surajkumar.mahato@tcs.com

## 1 Use Case Summary Table

| Item | Details |
|---|---|
| Category | 5G |
| Problem Addressed | Clearly describe the primary issue or challenge |
| Key Aspects of Solution | 1) 3D remote education experience with very low bandwidth: It enables trainees or students to interact with a live 3D virtual representation of a distant teacher/demonstrator in real-time with a very low-bandwidth consumption. The bandwidth savings is very significant compared with 3D and even with 2D visual transmission which is in the order of several Gbps and Mbps respectively.<br>2) Democratized scalable solution with no need for specialized infrastructure: Unlike holographic 3D video transfer, it does not require any specialized sensor, camera and studio set up at both end points. The teacher/ demonstrator only needs a RGB camera attached to a computer. This solution can be in generalized as cost-effective, scalable 3D telepresence with acceptable realism.<br>3) AI-native Semantic Communication with optional GenAI integration: It uses semantic communication to predict the body pose and position of the distant teacher / demonstrator in real-time using AI and transmits the predicted information to the trainee / student-end for mimicking a pre-stored 3D body model. It uses artificial intelligence to predict the body posture of the distant teacher in real-17th European Conference on Technology Enhanced Learningtime, and the figure is displayed in situ using augmented reality. GenAI may be used to virtually recreate some of the objects associated with the teacher, at the trainee/ student's end. |
| Technology Keywords | Semantic Communication, human pose estimation, natural language processing, augmented reality. |
| Data Availability | For body modelling we used the datasets used in [2]. For motion prediction we used [3]. |

(continued)

| Item | Details |
|---|---|
| Metadata (Type of Data) | Visual |
| Model Training and Fine-Tuning | In the initialization phase a 3D body model is initialized through Parametric Human Model Creation utilizing the SMPL-X [2] parametric human model. To create the SMPL-X model from image sequences, we use the method in [11]. The initial body model is shared with the receiving end before transmission. For human motion capture we extract temporally consistent 3D human pose and shape from monocular video through enhanced spatio-temporal context by extracting body-aware deep features [3] from individual frames and simultaneously predicting initial per-frame estimates of body pose, shape, and camera pose using a standard method [5]. The final motion capture output is presented in a JSON format and transmitted for rendering at the other end. |
| Testbeds or Pilot Deployments | Initial PoC and the related experience is published as referred in [1] in the references section. |
| Code repositories | N/A |

# 2 Use Case Description

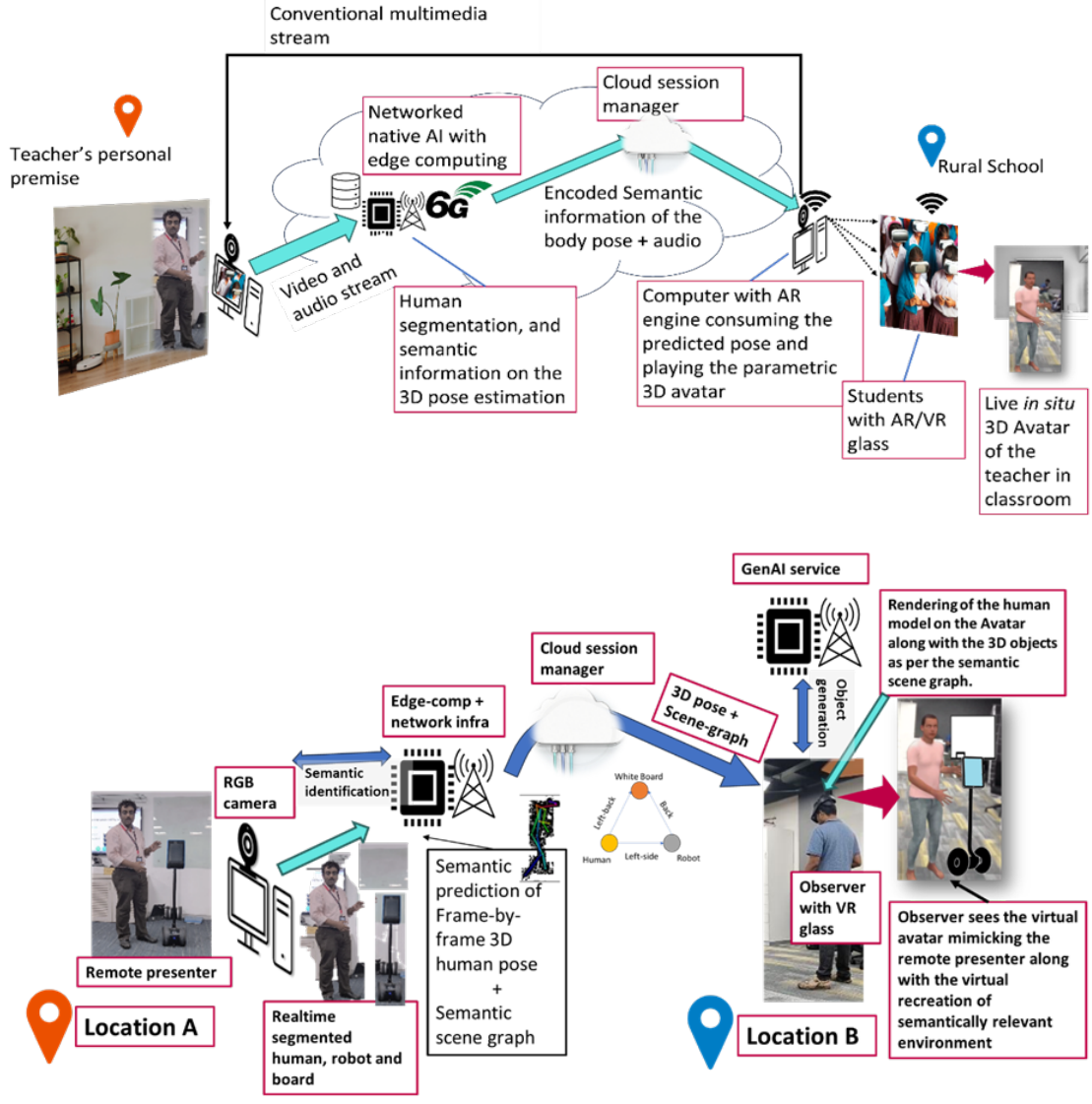
## 2.1 Description

Lack of good teachers/ expert trainers is one of the hindrances against quality education and training for many developing and underdeveloped nations [5]. The present use case addresses this gap using AI native semantic communications. Although remote teaching using telepresence robots has been considered, it suffers from infrastructural cost and challenges [6], lacks humane touch and does not enable the exchange of non-verbal body cues like realistic communication. So, there is a shift in demand for 3D telepresence [7] with an expectation to bring more realism into teacher-student interaction. 3D holoportation, as described in [8], is not scalable, requires huge bandwidth, and prohibitively costly for any democratized usage due to dedicated infrastructure requirements. So, inspired by the future looking education roadmap in countries like India embracing augmented reality (AR) and virtual reality (VR) glasses [9], we propose the following solution as described by remote teaching scenario.

As shown in Fig. 1, the teacher stands in front of a simple RGB camera attached to a computer which streams live visuals and audio of the teacher to an edge computer natively integrated with the 5G/6G network service. The edge computer is equipped with an AI algorithm which extracts the teacher's 3D body posture in real-time at a specified framerate. The extracted posture is encoded as semantic information of the body pose and transmitted over the network to the remote school's computing device, which is connected to the VR glasses of the students. The computing device at the school has a local AR engine which is preloaded with a parametric 3D avatar of the teacher. The semantic information of the live body posture received by the computing device in the school is decoded and transferred to the 3D avatar. As an effect, each student sees the remote teacher's avatar in situ as if the teacher is present in the classroom. We call this 'Semantic Live Streaming'. Since the teacher does not need to see the students in 3D, a camera attached to the computing device in the school can transmit back the view of the class to the teacher through conventional real-time streaming. The teacher does not need a

net specialized set or h/w unlike the systems that require special body-sensors to be attached to get the body pose [10]. The actual transmission over the network backhaul takes only about a few hundred bytes to a few kilo bytes per frame. This is much less than 3D video streaming or even 2D video. The solution may also be deployed for rendering on digital video screen if AR/VR glass and associated infrastructure is not available.

As a further enhancement, GenAI service may be added to the receiving end to replicate the necessary objects relevant to the remote trainer. As shown in Fig. 2, the remote trainer is demonstrating a robot which does not exist at the trainee's end (observer in the figure). The transmit-side intelligence may generate and send a scene description along with the 3D pose information. The GenAI service understands the objects and the relations from the scene description and recreate those objects.




**Use Case Status:**

An initial PoC to understand the feasibility has been done and the relevant experiences have been reported in the publication in [1].

## 2.2 Benefits of the Use Case

This solution addresses the shortage of expert teaching staff by enabling immersive, remote knowledge-sharing experiences that can be delivered with minimal infrastructure. It facilitates realistic and interactive live teaching across geographically remote or rural areas, enhancing student engagement and improving access to quality education. By narrowing the gap in educational opportunities across different regions, the use case contributes to greater equity in learning outcomes.

Additionally, the platform offers strong potential for transforming professional training and upskilling programs. It enables trainers to deliver scalable, high-quality sessions remotely, fostering equal opportunities for workforce development and expanding access to economic growth. By supporting wider reach and efficiency in knowledge delivery, the use case helps create a more inclusive and adaptable training ecosystem.

## 2.3 Future Work

Following actions are planned:

- The initial implementation lacks the desired real-time performance. It has been observed that we need to do more improvement in the human motion prediction algorithm. Some algorithmic enhancement is planned.
- At present the semantic encoding part is happening at the endpoint itself. This needs to be transferred to an edge computing facility. This will add some more complexity as the video needs to be transmitted to the Edge infrastructure. We are considering whether splitting the AI operation between the Edge and the endpoint may help.
- A standardization of the data format for the semantic exchange is to be devised for transferring the human motion parameters.
- The application-layer protocol is to be designed with proper loss-resilience and recovery mechanisms.
- The audio of the teacher is to be integrated with the semantic body description and a standard to be proposed.

Additional resources:

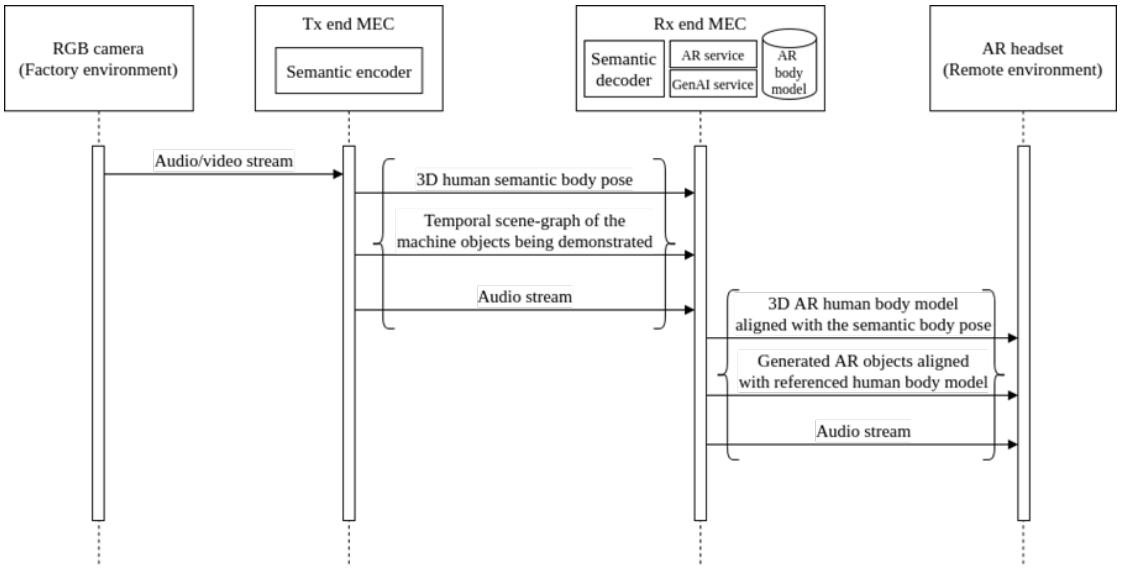
- The human motion prediction model needs to improve for more accurate prediction and a new data set to be created.
- The network infrastructure with the Edge service to be made available.

# 3 Use Case Requirements

- **REQ-01**: It is critical that the Edge computing services with sufficient GPU are available near the endpoints along with the network.
- **REQ-02**: It is critical that the Edge computing infrastructure must allow the application service provider to host the necessary AI algorithms as specified.
- **REQ-03**: It is critical that a cloud-service must be available to maintain the session.
- **REQ-04**: It is critical that the algorithms for generating the semantic information related to human motion-capture perform in real-time (approximately in 100ms considering a framerate of 10 fps)
- **REQ-05**: It is of added value that the rendering algorithm is adaptive considering whether the trainees are watching through AR/VR glass or watching through large display screens.

- **REQ-05**: It is of added value that the rendering engine in equipped with GenAI algorithms to create small objects as per scene descriptions and can attach to at the proper scene locations using AR.
- **REQ-07**: It is expected that the semantic information is multiplexed with audio in a synchronized manner.
- **REQ-08**: It is expected that a parametric human body model must be available at the initial stage and shared with the transmitter.

## 4 Sequence Diagram




## 5 References

[1] A. Bhattacharyya, A. Sau and S. Mahato, "Semantic Communication with Networked AI in 6G: Application Scenarios, Insights from Practical Experiments, Standardization Aspects," 2025 17th International Conference on COMmunication Systems and NETworks (COMSNETS), Bengaluru, India, 2025, pp. 1264-1268, doi: 10.1109/COMSNETS63942.2025.10885668.

[2] G. Pavlakos, V. Choutas, N. Ghorbani, T. Bolkart, A. A. A. Osman, D. Tzionas, and M. J. Black, "Expressive body capture: 3D hands, face, and body from a single image", IEEE Conf. on Computer Vision and Pattern Recognition (CVPR), 2019.

[3] C. Zhongang, et. al., "Humman: Multi-modal 4d human dataset for versatile sensing and modeling", European Conference on Computer Vision, 2022.

[4] N. Natalia, D. Novotný, V. Khalidov, M. Szafraniec, P. Labatut and A. Vedaldi, "Continuous Surface Embeddings", ArXiv abs/2011.12438, 2020.

[5] CRY, "Education In Rural India: Challenges, Opportunities, Initiatives", May 31, 2024, URL: https://www.cry.org/blog/education-in-rural-india/ .

[6] M. Perifanou, A. A. Economides, P. Häfner, T. Wernbacher, "Mobile Telepresence Robots in Education: Strengths, Opportunities, Weaknesses, and Challenges", Educating for a New Future: Making Sense of Technology-Enhanced Learning Adoption (EC-TEL), 2022.

[7] Market Digits, "Immersive Connections: Exploring the 3D Telepresence Market and its Transformative Impact", January, 2024. URL: https://www.newstrail.com/immersive-connections-exploring-the-3d-telepresence-market-and-its-transformative-impact/ .

[8] S. Orts-Escolano, et. al, "Holoportation: Virtual 3D Teleportation in Real-time", 29th Annual Symposium on User Interface Software and Technology (UIST '16), 2016.

[9] K. Swargiary, "The Future of Education in India: A Comprehensive Study on the Integration of Virtual Reality (VR) Technology in Schooling", SSRN Electronic Journal, 2023.

[10] "Sony takes a deep dive into the Metaverse with motion-tracking wearables", URL: https://www.forbesindia.com/article/cryptocurrency/sony-takes-a-deep-dive-into-the-metaverse-with-motiontracking-wearables/81629/1 . Last accessed: 12th Feb., 2024.

[11] Y. Xiu, J. Yang, X. Cao, D. Tzionas, and M. J. Black, "ECON: Explicit Clothed humans Optimized via Normal integration", IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June, 2023.

# Use Case 20: AI-Driven Free Space Optical Communication System

**Organization:** Mbeya University of Science and Technology (MUST)

**Country:** Tanzania

**Contact Person(s)**:

Silvester Beatus Mhagama, Email: silvestersilvester36362@gmail.com

Dr. Joseph Sospeter, Email: joseph.sospeter@must.ac.tz

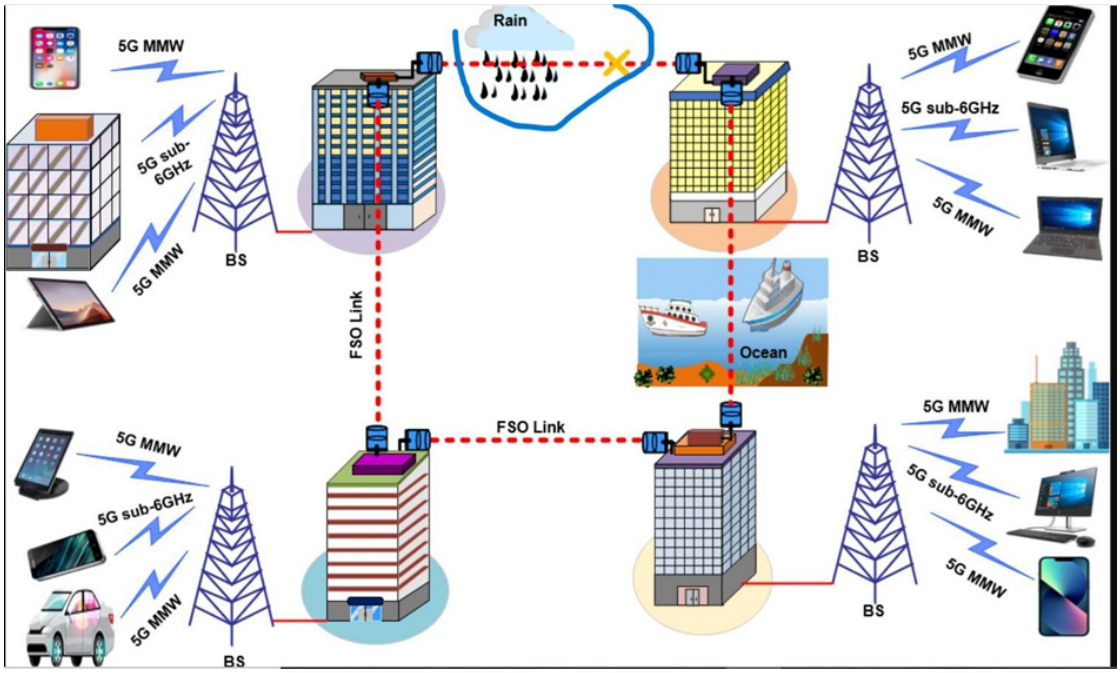
## 1 Use Case Summary Table

| Item | Details |
|---|---|
| Category | 5G |
| Problem Addressed | Severe signal degradation in Free-Space Optical (FSO) communication due to weather conditions (fog, rain, turbulence). |
| Key Aspects of Solution | AI-driven dynamic power control, hybrid FSO/RF switching, real-time weather data integration, adaptive modulation |
| Technology Keywords | AI/ML, Optical Communication,<br>Atmospheric Modeling, IoT Integration, 5G/6G |
| Data Availability | Public: TMA weather datasets.<br>Private: Simulation data from MUST labs. |
| Metadata (Type of Data) | Time-series visibility, rain rate, humidity; signal strength (dBm), BER logs. |
| Model Training and Fine-Tuning | LSTM networks for weather prediction; reinforcement learning for adaptive power optimization.<br>The AI component uses time series models to predict optical signal attenuation over time. We are currently experimenting with LSTM (Long Short-Term Memory) and Prophet models to forecast the signal strength based on environmental variables |
| Testbeds or Pilot Deployments | MATLAB Simulink prototype.<br>We verify the AI predictions against simulated optical signal data in MATLAB Simulink, using the actual weather conditions as input. Metrics such as RMSE (Root Mean Square Error) and BER (Bit Error Rate) were used to validate accuracy. |
| Code repositories | AI-driven FSO/mwanasheria/GitHub available on GitHub |

# 2 Use Case Description

## 2.1 Description

Free Space Optical (FSO) communication offers a promising high-speed alternative to traditional radio frequency (RF) systems, particularly in environments with high bandwidth demand. However, the reliability of FSO links is significantly impacted by environmental factors such as fog, rain, turbulence, and physical vibrations, which can lead to signal degradation and communication loss.




This use case proposes an AI-driven FSO system designed to maintain reliable high-speed communication by dynamically adjusting transmission parameters in real-time. By integrating artificial intelligence techniques, the system addresses common atmospheric disruptions and automates key aspects of link management, such as beam alignment and modulation adaptation.

The core objectives of this solution are to improve FSO link reliability under changing atmospheric conditions, automate beam alignment using AI, and dynamically adapt modulation and error correction schemes to maintain optimal signal integrity. The system leverages predictive AI models for atmospheric compensation, reinforcement learning for adaptive modulation, and real-time optimization of bit error rate (BER) and overall link quality.

The system behaves like a self-aware communication link – constantly analyzing its environment and optimizing performance through intelligent adaptation. It maintains high data integrity and link reliability even in the presence of weather disruptions or physical vibrations by:

- Predicting and reacting to atmospheric conditions,
- Dynamically tuning optical parameters, and
- Learning optimal responses from past link behavior.

This use case marks the first-time integration of AI for real-time modulation and environmental adaptation within FSO systems. It reduces pressure on congested RF bands and lays the foundation for resilient optical networks. The system is also energy-efficient, produces no RF noise, and is adaptable to evolving climate conditions. Its scalability makes it well-suited for a variety of applications, including smart cities, disaster zones, remote education, and even space-based communications.

**Use Case Status:** The design and initial simulation of the FSO system are complete in MATLAB Simulink. The AI model for predicting signal attenuation is currently in the training phase using historical climate datasets. We expect to begin integration and testing within the next two weeks.

**Partners** : Mbeya university of science and technology (Supervision and technical support)

Tanzania Metrological Authority(TMA) For Weather Datasets

## 2.2 Benefits of the use case

- This use case promotes innovation in telecommunications infrastructure by applying AI-driven solutions to enhance the reliability and performance of Free Space Optical communication systems.
- It also supports sustainable smart city communications by enabling high-speed, congestion-free data transmission, which is crucial for connected urban services and emergency response networks.
- Additionally, the system adapts communication strategies in response to climate-related challenges such as atmospheric disruptions, helping ensure consistent and resilient connectivity under varying environmental conditions.

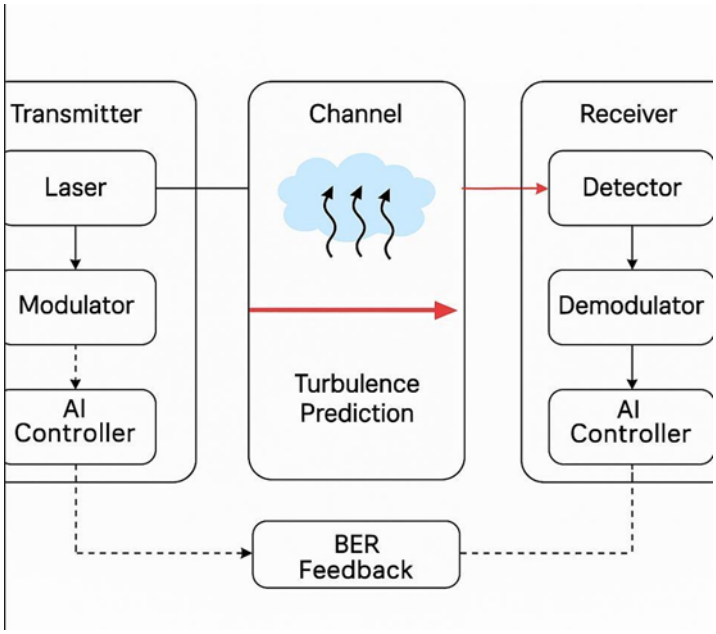
## 2.3 Future Work

- Hardware prototype implementation.
- Field testing under varied climate.
- Deployment with 6G and LEO satellites.
- Federated AI for decentralized optical networks.

## 3 Use Case Requirements

- **REQ-01**: It is critical that the system supports AI-assisted optical beam tracking and alignment to maintain communication stability under environmental disruptions.
- REQ-02: It is expected that the solution enables real-time adaptive modulation switching to dynamically respond to changing transmission conditions.
- **REQ-03**: It is critical that atmospheric turbulence models are integrated as input to optimize system performance and predict signal degradation scenarios.
- **REQ-04**: It is expected that the optical hardware and receivers are embedded with AI modules to enable intelligent, on-device processing and decision-making.
- **REQ-05**: It is of added value that the system includes a simulation and testing platform, such as MATLAB Simulink, to validate performance under various operational scenarios.

## 4 Sequence Diagram




## 5 References

[1] Bahaa E. A. Saleh and Malvin Carl Teich, Fundamentals of Photonics, 3rd ed. (Hoboken, NJ: Wiley, 2019).

[2] International Telecommunication Union Telecommunication Standardization Sector (ITU-T), ITU-T FSO Recommendations (Geneva: ITU-T).

[3] "IEEE Papers on ML in Optical Systems," IEEE Xplore.

[4] "TMA Datasets for Atmospheric Analysis," [Publisher/Organization responsible for TMA datasets].

[5] National Aeronautics and Space Administration (NASA), "NASA Earthdata," accessed June 24, 2025, https://earthdata.nasa.gov/.

[6] National Oceanic and Atmospheric Administration (NOAA), "NOAA Climate Data," National Centers for Environmental Information (NCEI), accessed June 24, 2025, https://www.ncei.noaa.gov/.

[7] MathWorks, "MATLAB Documentation for Optical Systems," MATLAB.

# 4.4. Category 4: Productivity

UNITED ARAB EMIRATES
MINISTRY OF COMMUNITY EMPOWERMENT

## Use Case 1: Non-Profit Organizations – AI Powered Registration Process

**Organization:** Ministry of Community Empowerment

**Country:** United Arab Emirates

Contact Person:

Abdulkader.burghul@moce.gov.ae

## 1 Use Case Summary Table

| Category | Community Empowerment |
|---|---|
| Problem Statement | Manual NPO registration causes delays due to complex validation |
| Solution Summary | AI engine validates NPO names in English and Arabic, checking for duplication and compliance |
| Technology Keywords | Natural Language Processing (NLP), Elasticsearch |
| Data Availability | Private |
| Metadata (type of data) | Text (background-text, prompt-text, response-text, response-correction) |
| Model Training and fine-tuning | Sentence-BERT fine-tuned for bilingual data (Arabic and English) |
| Testbeds or Pilot Deployments | The pilot deployments included scenarios for duplication detection, prohibited term filtering, and generation of alternative names in both English and Arabic. |

## 2 Use Case Description

### 2.1 Description

The United Arab Emirates (UAE) is implementing new platforms and procedures for non-profit organizations to enhance their regulation, support, and integration within the social sector. This initiative aims to empower these entities to operate effectively within a clearly defined legal framework, ensuring compliance with applicable laws and regulations. By doing so, the UAE

seeks to bolster the contributions of non-profit organizations in addressing societal challenges and improving the quality of life for its communities.

The AI engine will validate Non-Profit Organization (NPO) names provided in both English and Arabic against specific criteria, including duplication, prohibited words, and generating alternative suggestions. The validation will leverage advanced Natural Language Processing (NLP) models to ensure accurate results [1].

*Environment Requirements and Deployment Process:*

The API is deployed and running via Uvicorn inside a configured Anaconda environment, and can be optionally registered as a Windows service (using NSSM).

Logging is enabled to collect runtime feedback.

Based on debugging logs and UAT testing, the system successfully:

- Connects to the SQL database
- Returns similar names
- Flags duplicates based on logic
- Returns appropriate validation messages

*Key Features:*

1) Duplication Check:

- Verify if the provided NPO name exists in the database in either language (English or Arabic).
- Notify the applicant with descriptive validation messages if:
  - The English name is already taken.
  - The Arabic name is already taken.
  - Both names are taken.
- Provide a link to the NPO National Registry for public information on existing NPOs.

2) Prohibited Words Check:

- Validate names against a list of prohibited terms (e.g., names linked to terrorism or culturally sensitive language).
- Prevent the user from saving prohibited names and display a descriptive message indicating the issue.

3) Alternative Name Suggestions:

- Generate alternative name recommendations using an AI model trained on linguistic patterns.
- Ensure suggested names comply with business rules.
- Enable the user to click on a [Generate Alternative Names] button to view the recommendations.

The validations are guided by predefined policies which are dynamically updated:

- Exact Match Policy: Disallows submission if the name exists exactly (either in English or Arabic).
- Similar Name Notification Policy: Flags and informs the user if a semantically or visually similar name exists.
- Prohibited Terms Policy: Names associated with terrorist organizations, government bodies, or restricted terms are blocked.

## 2.2 Benefits of the use case

Industry, Innovation, and Infrastructure

- The NPO platform provides support to startups, research institutions, and social enterprises driving technological and social innovation.
- It enables infrastructure-focused NPOs to operate effectively in underserved areas—supporting projects like schools, hospitals, clean water, and sustainable transport.

Partnerships for Sustainable Development

- The platform connects NGOs, governments, and private and international actors, enabling joint development initiatives.
- It supports knowledge sharing, coordination, and project alignment through centralized digital infrastructure.

Promoting Data Sharing and Transparency

- Open data features allow stakeholders to track development outcomes and inform decision-making.
- AI and analytics help identify gaps and optimize resource distribution.

## 2.3 Future Work

The next development phase will expand the AI engine to include volunteer opportunity matching. Key enhancements will include:

- Skill-Based Matching: Volunteers will be paired with projects aligned with their expertise.
- Geolocation & Availability: Opportunities will be suggested based on proximity and time availability.
- Behavioral Insights: AI will analyze previous engagement to improve matching accuracy.

Additionally, AI-driven retention strategies will be implemented:

- Engagement Prediction: Identifying volunteers at risk of disengagement.
- Leadership Identification: Recommending leadership roles for highly engaged volunteers.
- Personalized Communication: Sending tailored reminders and recognition messages.

Example: A volunteer who frequently participates in food distribution may receive a message like: "You've supported 500 families this year—would you like to join our leadership program?"

## 3 Use Case Requirements

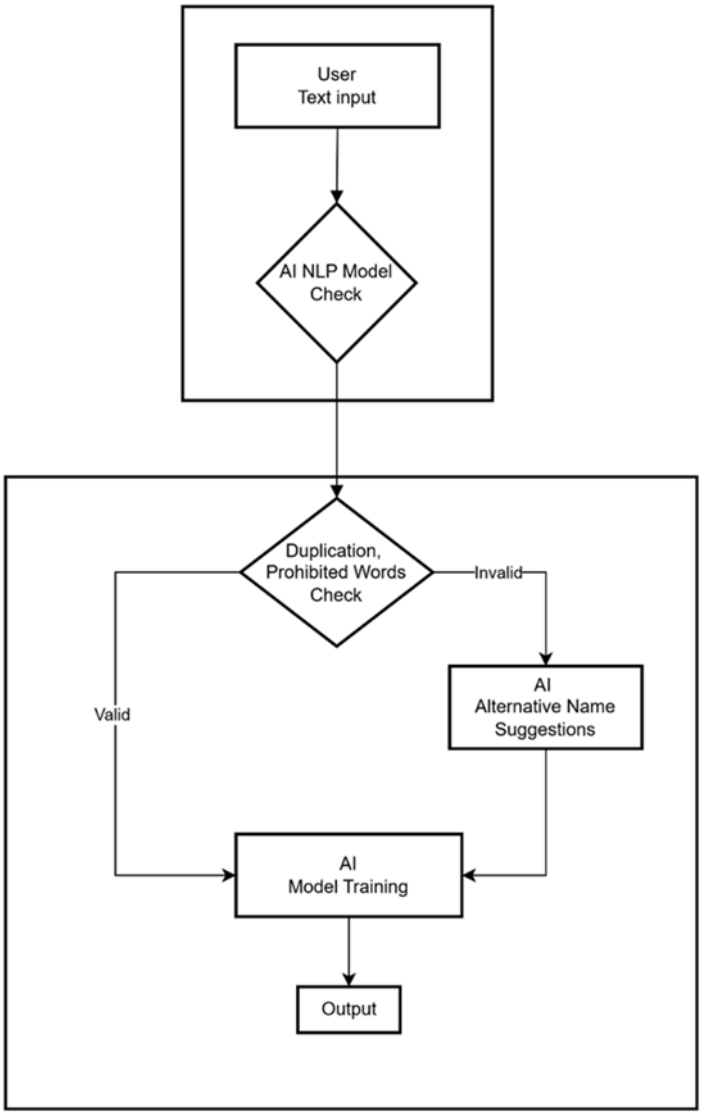
- **REQ-01**: Machine Learning Models – predictive analytics, analyzing historical data to promote NPO initiatives and volunteers matching / retention
- **REQ-02**: User friendly interface – To interact with the AI driven name validator

AI Models of the solution

The endpoint relies on Natural Language Processing (NLP) models, specifically:

- Semantic Similarity Models (Sentence-BERT [1]) to compare the meaning of names in Arabic and English.
- Fuzzy Matching Algorithms ( Levenshtein distance, token sort ratio) for partial or character-level similarity.
- These models act like a recommendation engine, suggesting close or semantically similar names based on a given input.

## 4 Sequence Diagram



Solution's Pipelines

Validation Pipeline (with Elasticsearch)

1) Input Collection

- User provides:
- English name
- Arabic name

2) Text Normalization

- Convert to lowercase
- Strip whitespace and special characters
- Normalize Arabic text (e.g., remove diacritics, standardize characters)

3) Exact Match Validation

- Check for exact matches in existing names (from SQL or Elasticsearch)
- If an exact match is found, flag it as not valid and return related record IDs

4) Semantic Similarity Search

- Convert input name into a semantic vector using multilingual Sentence-BERT [1]
- Query Elasticsearch's vector index [2] using cosine similarity
- Retrieve top N semantically similar names

5) Prohibited Words Check

- Validate against a list of banned or sensitive words (e.g., political groups, terrorist organizations)
- Reject name if it contains any prohibited terms

6) Policy-Based Validation

Apply name validation rules, such as:

- Domain-specific checks: Flag names that closely resemble ministry functions or other restricted domains

7) Decision Logic & Thresholding

- If:
- No exact match
- No prohibited words
- Semantic similarity score is below defined threshold
- Name complies with all policies – Then name is valid

8) Response Generation: Return a structured response

Name Recommendation Pipeline

1) Input Capture

- Input name
- Language
- Reason for rejection

2) Text Cleaning & Normalization

- Remove legal terms, trim spaces, standardize text
- Normalize Arabic: remove diacritics, normalize alef/hamza, etc.

3) Rule-Based Variants

- Apply rule-based methods as before: synonym replacement, reordering, suffixes
- Generate a base pool of candidate names

4) Sentence-BERT Embedding

- Use multilingual SBERT model like paraphrase-multilingual-MiniLM-L12-v2 [3]
- Convert the input name to a dense embedding
- Precompute embeddings for existing names in the database and store them in:
- Elasticsearch vector index (with dense_vector)

5) Semantic Search

- Use cosine similarity between input name and stored name embeddings
- Retrieve top N semantically similar names

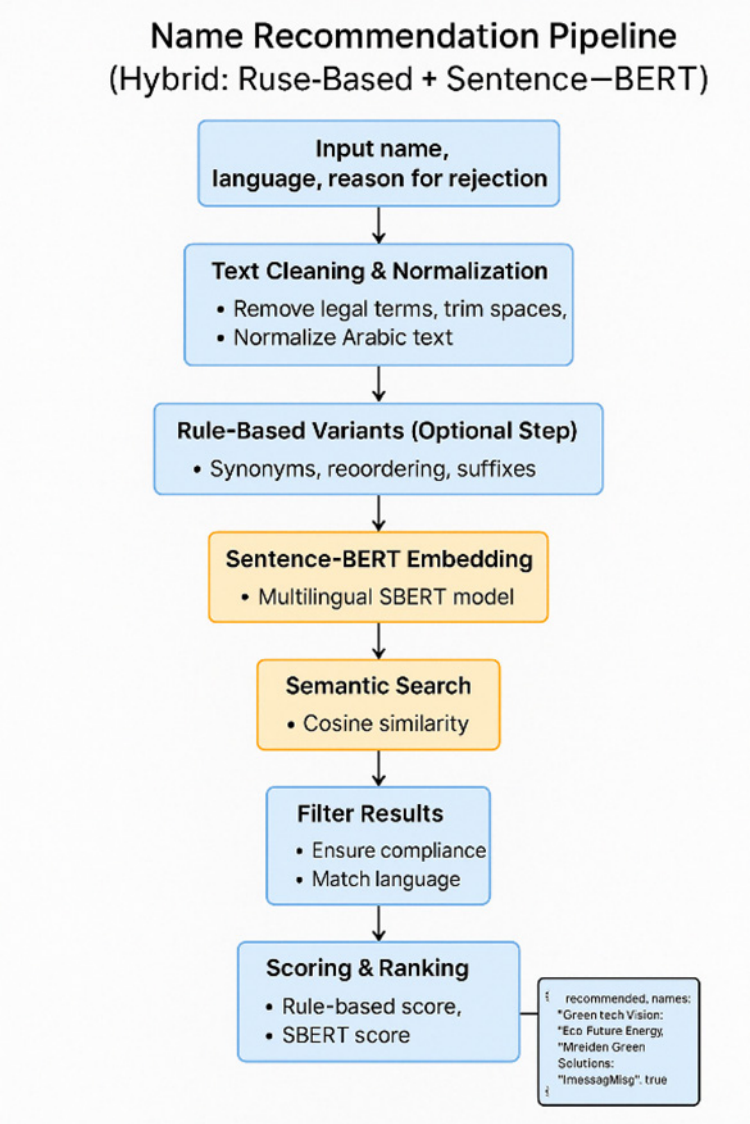
6) Filter Results

- Remove any names that caused rejection (e.g., exact match or prohibited)
- Filter based on policy: sensitive words, reserved names, etc.
- Prioritize names that:
- Are semantically similar but legally and contextually acceptable
- Match user's language preference

7) Scoring & Ranking

- Combine scores:
- Rule-based score (linguistic similarity, uniqueness, policy compliance)
- SBERT score (semantic similarity)
- Weighted average or customizable scoring strategy

8) Response Generation: Return a structured response




## 5 References

[1] Reimers, N., & Gurevych, I. (2019). Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks. arXiv:1908.10084.

[2] Elasticsearch. https://www.elastic.co/elasticsearch/

[3] paraphrase-multilingual-MiniLM-L12-v2. https://huggingface.co/sentence-transformers/paraphrase-multilingual-MiniLM-L12-v2

# Use Case 2: AI-Powered GaP Repository: Democratizing Access to Political Knowledge for Young Women in South Asia

**Organization:** Centre for Gender and Politics

**Country**: India

**Contact Person(s)**:

Akhil Neelam, akhil@cgapsouthasia.org

Sugandha Singh Parmar, contact@cgapsouthasia.org

## 1 Use Case Summary Table

| Item | Details |
|---|---|
| Category | Productivity |
| Problem Addressed | Young women in South Asia face limited access to personalized political knowledge, mentorship opportunities, and language-accessible resources, resulting in underrepresentation in politics. |
| Technology Keywords | NLP, Multilingual, Machine Learning, Knowledge Graph, Conversational AI, |
| Key Aspects of Solution | - AI-driven political information delivery (contextualized)<br>- Multi-lingual Natural Language Processing (NLP) support<br>- AI mentorship matching<br>- Interactive learning modules |
| Data Availability | Private (Existing CGAP resource database) + curated public data on gender and politics |
| Metadata (Type of Data) | Text (articles, reports), Structured data (candidate databases), Conversational data (queries) |
| Model Training and Fine-Tuning | N/A |
| Testbeds or Pilot Deployments | https://www.cgapsouthasia.org/gap-repository<br>In our first phase, we launched a Content Management System (CMS) based repository of resources for young women in politics. |

# 2 Use Case Description

## 2.1 Description

The AI-Powered GaP Repository transforms the existing Centre for Gender and Politics static resource database into an interactive conversational platform designed to empower young women entering South Asian politics. Despite representing half the population, women remain significantly underrepresented in regional political institutions - in India, only 14% of parliamentarians are women, with just 14 under age 40 in a 543-member parliament.

Traditional political education approaches face critical limitations: written resources lack personalization, existing mentorship programs have limited reach, and language barriers persist across diverse linguistic communities. Young women aspiring to political careers struggle to access relevant information, find suitable mentors, and navigate complex regional political landscapes.

Our AI solution leverages Natural Language Processing and machine learning to provide personalized guidance through four key functions: contextual political information delivery, multi-lingual support for South Asian languages, AI-facilitated mentorship matching, and interactive learning modules tailored to users' experience levels.

The AI-based approach offers significant advantages: 24/7 accessibility regardless of location, personalized guidance based on individual contexts, scalable mentorship beyond traditional program limitations, and multi-lingual support that democratizes access across linguistic communities. The platform continuously learns from user interactions to improve recommendations.

Implementation challenges include ensuring data privacy and security, addressing technical limitations in areas with poor connectivity, and maintaining relevance in rapidly changing political landscapes. Despite these challenges, the AI-powered GaP Repository represents a transformative tool to increase women's political participation across South Asia, creating more representative and effective democratic institutions.

**Partners**

- For the current version of our GaP Repository, we worked with JPMorgan&Chase as part of their Force For Good program.
- For our upcoming AI-powered GaP Repository version, we are yet to finalize the partners. Tentatively, we will be engaging with technology institutes (colleges, universities) and nonprofits that work with women in politics.

## 2.2 Benefits of the Use Case

By breaking down informational barriers and providing guidance on navigating male-dominated political structures; the platform helps overcome systemic barriers that have historically excluded women from political power.

The platform's multi-lingual capabilities ensure language is not a barrier to political education, particularly benefiting women from marginalized communities who face intersectional challenges.

Research shows women politicians often prioritize education, healthcare, and addressing gender-based violence, strengthening democratic institutions.

## 2.3 Future Work

Our implementation strategy for the AI-powered GaP Repository follows a four-phase approach to transform CGAP's existing static resource database into a dynamic, AI-driven platform that addresses the unique challenges facing young women in South Asian politics.

Phase 1 focuses on data enrichment and AI model development. We'll convert CGAP's existing repository content into training data, supplemented with region-specific political information from verified sources. This will be structured into a knowledge graph that captures the nuances of political systems across South Asian countries. Simultaneously, we'll develop and fine-tune a multilingual large language model specialized for political discourse in regional languages including Hindi, Urdu, Bengali, Tamil, and Sinhalese.

Phase 2 centers on building the conversational interface and personalization system. The platform will feature a progressive profiling system that learns user preferences and career stage without intrusive questioning. We'll develop contextual recommendation algorithms that deliver tailored resources based on the user's specific political environment, from local governance to national politics.

Phase 3 implements the mentorship matching feature, utilizing an AI system that analyzes mentor profiles and user needs to suggest optimal connections. This will incorporate availability scheduling and conversation prompts to facilitate meaningful engagement.

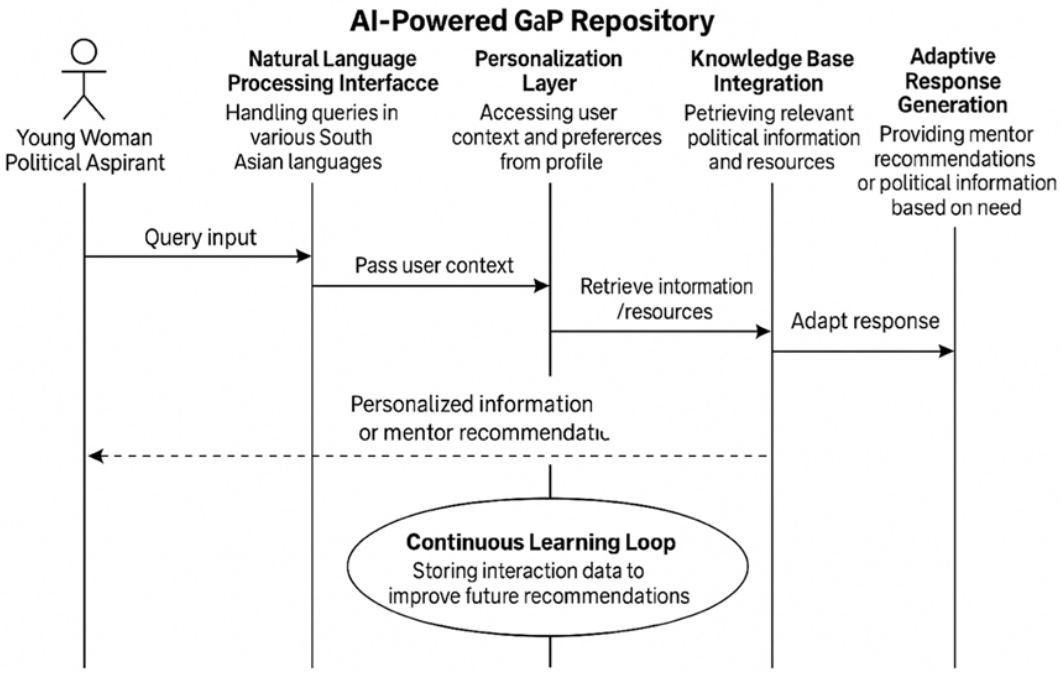
The final phase focuses on user testing and feedback integration, involving young women from diverse backgrounds across South Asia. We'll implement a continuous improvement system that learns from user interactions to enhance recommendations and content relevance. We intend to give controlled public access for NGOs, universities to get feedback and then scale the platform.

Throughout development, we'll prioritize data privacy and security through end-to-end encryption and anonymized data processing. For areas with connectivity challenges, we'll develop offline functionality that syncs when connections become available. The platform will be designed for scalability, allowing for expansion to additional regions while maintaining contextual relevance to local political realities.

# 3 Use Case Requirements

- **REQ-01**: Multilingual NLP support (Hindi, Urdu, Bengali, Tamil)
- **REQ-02**: Secure user data storage and anonymized learning logs
- **REQ-03**: Dynamic recommendation system based on user profile + region
- **REQ-04**: Mentor database with matching algorithm
- **REQ-05**: Offline functionality for low-connectivity regions
- **REQ-06**: Scenario-specific modules (e.g., filing nominations during elections)

## 4 Sequence Diagram




## 5 References

[1] UN Women reports on women's political participation

[2] CGAP's static repository (internal database) and research www.cgapsouthasia.org

[3] Technical resources on multilingual LLMs (e.g., HuggingFace, BigScience)

# Use Case 3: HEART Engine is an AI-Powered Public Interest Tech Certification Platform to Build Future AI Public Interest Workforce

**Organization**: #BlackTechFutures Research Institute

**Country**: United States

**Contact Person(s)**:

Dr. Fallon Wilson, fallon@blacktechfutures.com

Ayori Selassie, iayori@gmail.com

Isaac McCoy, imccoy@stillman.edu

## 1 Use Case Summary Table

| Item | Details |
|---|---|
| Category | Productivity |
| Problem Addressed | Artificial Intelligence is advancing faster than societies, workforces, and governments are prepared for worldwide. As AI reshapes economies, public services, and daily life, communities everywhere—especially in vulnerable, rural, and underserved regions—face the risk of being left behind without a fair opportunity to adapt or influence the future. Artists from historically disadvantaged communities are not properly cited for their work in AI Art Generators (Small, 2023), culturally-sounding names receive low pay options based on large language learning models (Haim, Salinas, and Nyarko, 2024), Black people are labeled as future criminals in comparison to White people (Angwin et al., 2016), and Black women are consistently mischaracterized or not recognized in facial recognition software (Buolamwini and Gebru, 1998; Najibi, 2020). At the same time, key decisions about AI systems are often made far from the people they affect, without transparency, accountability, or respect for local cultures, values, and needs (Sources). Without a trusted and skilled workforce of public-minded technologists to bridge AI innovations with community priorities, nations risk deepening inequality, eroding public trust, and losing control over how AI shapes their societies. We must equip people everywhere, not just those in major technology hubs, to lead, govern, and thrive in the emerging AI-driven world. |

(continued)

| Item | Details |
|---|---|
| Key Aspects of Solution | HEART Engine builds the workforce and tools needed to ensure that AI serves communities, not just corporations or centralized systems.<br>The solution includes:<br>• Adaptive Training Pathways Aligned to Real-World Needs: HEART Engine delivers personalized learning journeys, designed to meet individuals where they are and prepare them for AI-integrated roles that advance the public interest.<br>• Synthetic Profile Agents to Guide Learning and Matching: Using synthetic profiles enriched with community data and regional trends, HEART Engine dynamically adapts training and job-matching to each learner's background, goals, and the needs of the places they serve.<br>• Bias-Aware Workforce Matching and Placement: AI-driven recommendation systems connect certified Public Interest Technologists to relevant jobs, fellowships, or innovation roles, with built-in bias auditing and explainability to ensure fairness and transparency.<br>• Regional Knowledge Base Development: HEART Engine captures structured insights from workforce data, innovation ecosystems, and cultural context to build living regional Knowledge Bases that inform both community action and platform improvements over time.<br>• Open-Source Tools for Broader Impact: Select components, including learning templates, bias mitigation tools, and ecosystem mapping methods, will be open-sourced to empower others to enhance digital equity and ethical AI efforts worldwide. |
| Technology Keywords | Adaptive Learning Systems, Recommendation Engines, Synthetic User Profiles, Knowledge Base Development, Workforce Analytics, Generative AI, Skills-to-Employment Matching Algorithms, Labor Market Analytics, AI Readiness Index, Explainable AI, Ethical AI, Community-Centered AI Design, Responsible AI Governance, Digital Equity Mapping, Bias Detection, Fairness Auditing, Explainable AI Tools, Equity-Centered AI Design, Human-in-the-Loop Systems, Synthetic Data Generation, Culturally Responsive Machine Learning, Localized Language Models, Micro-Credentialing Platforms, AI-Driven Job Matching, Skills Gap Analysis, AI Readiness Assessment, Digital Equity Infrastructure Mapping |

(continued)

| Item | Details |
|---|---|
| Data Availability | HEART Engine integrates multiple types of data through a structured process designed to create a rich, hyperlocal knowledge foundation for training and workforce development:<br>Primary Data Collection:<br>• Website Content Scraping: HEART Engine systematically scrapes and analyzes publicly available information from regional nonprofit organizations, civic tech groups, STEM accelerators, and innovation hubs to build a real-time map of local ecosystems.<br>• Community Surveys: Targeted surveys are conducted within regional communities to gather residents' perspectives on digital equity, technology adoption, and workforce needs, ensuring direct input from underrepresented populations.<br>• Manual Aggregation of Cultural Artifacts: The HEART Engine manually curates essential regional cultural elements, including local music history, entertainment trends, regional vernacular, and significant cultural milestones, by analyzing public sources, local media archives, and historical records.<br>Secondary Public Databases:<br>HEART Engine pulls structured datasets from trusted public repositories, including:<br>• Government Labor Databases (e.g., U.S. Bureau of Labor Statistics, Census American Community Survey)<br>• Education Datasets (e.g., National Center for Education Statistics, Higher Education APIs)<br>• Digital Equity Reports (e.g., NTIA Digital Equity Act data, regional broadband access maps)<br>These sources provide macro-level context for local workforce trends, digital divides, and educational attainment.<br>AI-Enriched Synthetic Cultural Data:<br>• Using Generative AI and natural language processing tools, HEART Engine processes local cultural data, such as song lyrics from regional artists, local news media coverage, and community storytelling projects, to extract meaningful cultural themes, values, and trends.<br>• This synthetic data is transformed into training modules that improve the cultural fluency of Public Interest Tech Workers, making them more effective at engaging communities in public interest tech work. |

(continued)

| Item | Details |
|---|---|
| Metadata | HEART Engine captures two types of data critical for platform optimization, adaptive learning, and workforce placement:<br>User-Level Structured Data<br>• Demographic Profiles: Age, education, geographic region, career background.<br>• Educational Attainment and Certifications: Completed modules, earned micro-certifications, and university partnerships.<br>• Job Placement Outcomes: Application rates, hiring success, retention metrics post-placement.<br>• Community Cultural Fluency Levels: Completion and scoring of localized cultural fluency modules.<br>• Digital Skills Progression: Baseline assessments and skills growth tracking over the learning journey.<br>System Metadata<br>• Engagement Analytics: User interaction logs, time spent per module, knowledge base search patterns.<br>• Learning Activity Timestamps: Module access dates, completion dates, and session durations.<br>• Feedback Submissions: Structured inputs from users about course quality and AI agent interactions.<br>This dual-layered data model enables HEART Engine to personalize learning pathways, predict successful job placement, continuously refine the AI models, and contribute to the Public Interest Tech Knowledge Base. |
| Model Training and Fine-Tuning | Adaptive Learning Prompt Templates, Recommendation Engine, Synthetic User Profile Agents, Voice and Tone Fine Tuning, Regional Language and Vernacular Training.<br>HEART Engine adapts and fine-tunes AI-driven systems to deliver culturally responsive, community-centered workforce training and placement. Rather than retraining foundational AI models, HEART Engine focuses on tailoring behavior, content, and matching processes for real-world effectiveness.<br>Key adaptation and fine-tuning processes include:<br>• Adaptive Prompt Engineering for Learning Personalization: Tuning dynamic learning prompts and pathways based on learner skill profiles, regional context, and progression analytics.<br>• Recommendation Engine Configuration and Optimization: Training supervised recommendation algorithms on structured learner data and synthetic profiles to match users with appropriate training modules, micro-credentials, and job opportunities.<br>• Synthetic User Profile Simulation: Generating and adapting synthetic learner personas to simulate diverse trajectories and better calibrate system responses to varied real-world needs.<br>• Voice and Tone Localization: Configuring conversational AI agents to adjust voice, tone, and engagement strategies based on regional communication styles and cultural expectations.<br>• Bias Detection and Fairness Monitoring Layering: Integrating bias detection models and fairness auditing tools into recommendation outputs to ensure transparent, equitable matching and decision-making processes. |

(continued)

| Item | Details |
|---|---|
| Testbeds or Pilot Deployments | At Spelman College and Ebenezer Baptist Church in Atlanta, Georgia, in the United States of America |
| Code Repositories | In development, to be hosted via GitHub for open access |

# 2 Use Case Description

## 2.1 Description

HEART Engine is an AI-powered certification and placement platform designed to build the next generation of Public Interest Technologists – trusted professionals who help communities navigate, shape, and benefit from emerging AI systems.

Rather than training only technical skills, HEART Engine develops public interest leaders who combine AI literacy, cultural fluency, and community-centered innovation.

The platform integrates multiple AI technologies to create a full-stack workforce development solution, including:

- Adaptive Learning Systems that personalize training based on each user's skill level, regional context, and professional goals.
- Synthetic User Profile Agents that simulate diverse learning trajectories and help match participants to targeted skill pathways and job opportunities.
- Recommendation Engines are optimized through bias auditing to ensure equitable and transparent job matching and placement.
- Regional Knowledge Base Development, capturing hyperlocal workforce trends, community needs, and cultural data to refine learning and placement strategies continuously.
- AI-Driven Digital Equity Mapping, including tools like the AI Readiness Index, to help communities, policymakers, and institutions understand where and how to strengthen their AI capacity.
- Use Case Status: ACTIVE – HEART Engine is currently part of a larger national research and ecosystem-building initiative to prepare diverse regional workforces for the AI-driven economy.

### Partners

- Spelman College, Meharry Medical College, and Stillman College are post-secondary anchors.
- Blacks in Technology Foundation, Civic Hacker, and Code Crew – training partners.
- AI Newbies, Microsoft, Omidyar Network, and Kapor Center – funders and infrastructure partners.
- Black Churches 4 Digital Equity – outreach and deployment.

## 2.2 Benefits of the Use Case

- personalized, AI-powered digital literacy programs.

- workforce-matching for equitable employment.
- mapping Black Tech Hubs
- bias mitigation in hiring.
- cross-sector collaboration in digital equity.

## 2.3 Future Work

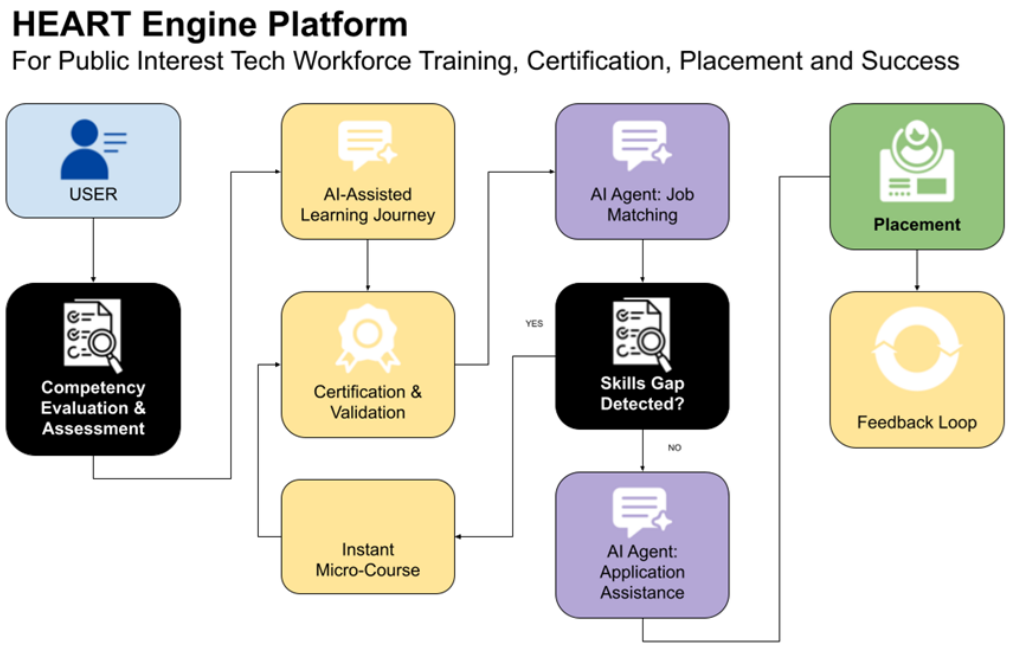
- Develop an open-source AI Public Interest Tech toolkit.
- Expand pilot cities beyond Atlanta.
- Scale partnerships with more HBCUs and faith institutions
- Integrate with the International Technology Union's ML standards and SMART infrastructure initiatives.

Develop additional simulation environments for training policymakers.

# 3 Use Case Requirements

- **REQ-01**: it is critical that digital literacy programs have Access to anonymized regional labor data.
- **REQ-02**: it is critical to Deploy adaptive AI learning modules.
- **REQ-03**: it is critical to achieve Interoperability with education and workforce APIs.
- **REQ-04**: it is critical to perform Bias auditing and explainability tools for hiring AI.
- **REQ-05**: it is expected to have GIS infrastructure data for innovative mapping tools.

# 4 Sequence Diagram



## 5 References

[1] Felicia Wilson, Dreaming of a Black Tech Future: Nashville City Report (2024).

[2] National Telecommunications and Information Administration, Digital Equity Act Overview (2023).

[3] Color of Change, Black Tech Agenda (2022).

[4] Kapor Center, Leaky Tech Pipeline (2018).

[5] International Telecommunication Union Telecommunication Standardization Sector (ITU-T), Recommendation Y.3172: Architectural Framework for Machine Learning in Future Networks (Geneva: ITU-T, 2019).

# Use Case 4: Making Everyone the Director of Their Own Life

**Organization**: PixVerse

**Country**: Singapore

**Contact Person(s)**:

Hu Zhengkun, huzhengkun@aishi.ai

Zhang Xiongyi, zhangxiongyi@aishi.ai

## 1 Use Case Summary Table

| Item | Details |
|---|---|
| Category | AI for Creativity and Digital Inclusion |
| Problem Addressed | Traditional video creation requires technical expertise and high-end equipment, excluding vast populations from participating in digital storytelling. |
| Key Aspects of Solution | No-code AI video generation, diverse templates, cultural adaptability, low hardware requirement |
| Technology Keywords | generative video, AI animation, democratized creation, multimodal AI, digital storytelling |
| Data Availability | Private; select results are publicly showcased on platforms like TikTok, YouTube |
| Metadata (Type of Data) | Prompts, Images, Videos, Templates, User-Generated Captions |
| Model Training and Fine-Tuning | Custom fine-tuning of diffusion models and video transformers using curated user behavior data |
| Testbeds or Pilot Deployments | https://pixverse.ai |

## 2 Use Case Description

### 2.1 Description

PixVerse is a generative AI platform redefining who gets to create in the digital age. Traditionally, video creation has been gated by technical expertise, expensive equipment, or social confidence. PixVerse breaks this barrier by offering intuitive, one-click video generation from simple inputs like a photo or a line of text – giving users cinematic results with no production knowledge required.

This aligns directly with **SDG 10: Reduced Inequalities**, by offering digital expression tools to anyone regardless of geography, education, income, or skill level. In many countries, less than 10% of social media users actually create videos, leaving the vast majority as passive consumers. PixVerse transforms this equation, inviting everyone to participate in the creator economy and amplify their voices.

PixVerse's real-world impact illustrates this inclusive potential:

- In Brazil, TikTok creator @nerdsoul used PixVerse to reimagine iconic characters through aging transformations—generating over 210 million views and gaining 300,000 followers in just two months. His achievements were covered by national television, highlighting the power of accessible AI creation tools.
- In Portugal, a Portuguese creator who spent decades in the fashion industry and reimagined his career with AI created a full 30-minute film entirely with PixVerse. The film was selected for the Fantasporto International Film Festival.

PixVerse has also sparked a bottom-up wave of everyday creativity. As of today, over 1.4 million videos have been published under the hashtag #pixverse on TikTok alone – many of them by first-time creators who had never posted videos before. This surge illustrates how intuitive AI tools can unlock new forms of digital self-expression at scale.

The platform offers a wide range of creative templates to serve diverse emotional and narrative needs:

- Sharing daily life – such as dancing pets and family moments;
- Self-enhancement – becoming stronger, cooler, or more stylish through muscle or beauty effects;
- Fan-based storytelling – recreating cinematic moments like superhero transformations;
- Emotional comfort – enabling imagined interactions like being hugged by Jesus or animating cherished old photos.

By offering such varied yet accessible tools, PixVerse truly makes creation a part of everyday life, bridging cultural, generational, and technological divides.

PixVerse also powers creative infrastructure for businesses and developers:

- Brands use PixVerse for cost-efficient ad creatives.
- Educators use it to transform lessons into animations.
- Developers embed its engine into broader workflows via SDKs and APIs.

Built on proprietary diffusion and transformer models, PixVerse enables keyframe control, motion generation, voice integration, and style transfer – all mobile-optimized. Its evolving platform ensures cultural adaptability and global scale.

UN Goals:

- SDG 10: Reduced Inequalities
- SDG 4: Quality Education
- SDG 9: Industry, Innovation and Infrastructure

Justification:

PixVerse supports SDG 10 by enabling inclusive storytelling tools. SDG 4 is supported through educational use cases, and SDG 9 through scalable infrastructure for creative industries.

Use Case Status: Operational

**Partners:** Independent creators globally; film festivals (e.g., FIRST International Film Festival)

## 2.2 Benefits of the use case

By offering such varied yet accessible tools, PixVerse truly makes creation a part of everyday life, bridging cultural, generational, and technological divides.

PixVerse also powers creative infrastructure for businesses and developers:

- Brands use PixVerse for cost-efficient ad creatives.
- Educators use it to transform lessons into animations.
- Developers embed its engine into broader workflows via SDKs and APIs.

Built on proprietary diffusion and transformer models, PixVerse enables keyframe control, motion generation, voice integration, and style transfer—all mobile-optimized. Its evolving platform ensures cultural adaptability and global scale.

## 2.3 Future Work

PixVerse aims to evolve into a universal AI video infrastructure across consumer, creator, and enterprise use cases:

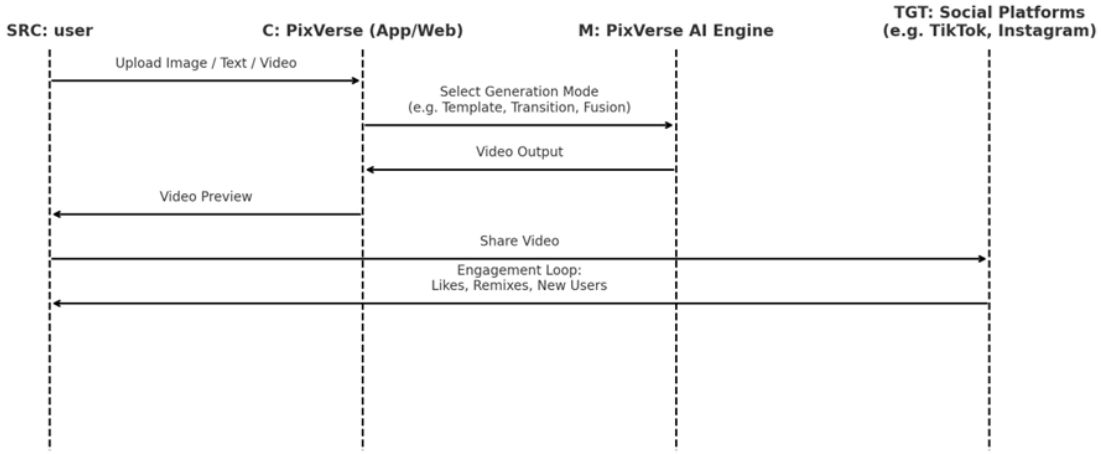
- Open SDK/API tools for developers and businesses.
- Improve motion realism via physics-based modeling.
- Expand localized template variety.
- Support real-time generation and mobile compatibility.
- Collaborate with nonprofits for inclusive storytelling.
- Release open datasets with ITU.
- Launch "Create with AI" global community program.

# 3 Use Case Requirements

- **REQ-01**: It is critical that user devices (mobile or web) support basic image input and basic GPU-accelerated processing for smooth real-time preview.
- **REQ-02**: It is critical that the AI models used are optimized for low-latency generation, to ensure wide accessibility across regions with low bandwidth or low-end devices.
- **REQ-03**: It is expected that localized content templates (cultural, language, regional styles) are available to encourage inclusive usage.
- **REQ-04**: It is expected that platform APIs and SDKs are made available for developers and third-party integrators to adopt the video generation engine.
- **REQ-05**: It is of added value if ethical filters and content safety modules are included to adapt the tool for sensitive educational or governmental use.

- **REQ-06**: It is of added value if anonymized usage data can be used to help fine-tune models for regional relevance and safety.

## 4 Sequence Diagram




## 5 References

[1] NerdSoul on TikTok. "Aging Spider-Man with AI", [Online]. Available: https://www.tiktok.com/@nerdsoul

[2] Rede TV! Brasil. "NerdSoul viraliza com vídeos de transformação com inteligência artificial", YouTube, May 2024. [Online]. Available: https://www.youtube.com/watch?v=htTyG0cNO0A

[3] Joao More. "AI Short Film Created Entirely with PixVerse", Vimeo, 2024. [Online]. Available: https://vimeo.com/1048952194

[4] Fantasporto Film Festival, "Official Selection 2024". [Online]. Available: https://fantasporto.com/pt-pt/movie-talks/

[5] TikTok, "#pixverse hashtag page", [Online]. Available: https://www.tiktok.com/tag/pixverse

# Use Case 5: Radiolytica - Leveraging AI Tools to Protect Communities from Harmful Content in Eastern DRC

**Organization:** Fondation Hirondelle

**Country:** RDC, Switzerland

**Contact Person**:

Gilles Magnin, gilles.magnin@hirondelle.org, +41797267594

## 1 Use Case Summary Table

| Item | Details |
|---|---|
| Category | Productivity |
| Problem Addressed | Radiolytica is an AI-driven system designed to monitor, transcribe, and analyze online community radio broadcasts in Eastern DRC, a region where radio remains the primary information source due to low internet penetration. The core problem addressed is the rampant spread of misinformation, disinformation, and hate speech through radio channels, which traditional manual monitoring methods cannot effectively mitigate. |
| Key Aspects of Solution | See initial application |
| Technology Keywords | AI, NLP, community radios |
| Data Availability | Public (community radio recordings) |
| Metadata (Type of Data) | Audio (community radio recordings) and text (transcriptions of recordings) |
| Model Training and Fine-Tuning | ON-the-shelf usage of industry leading tools (Speech-to-text, LLM) |
| Testbeds or Pilot Deployments | Not available |
| Code repositories | Not available |

# 2 Use Case Description

## 2.1 Description

Radiolytica is an AI-driven system designed to monitor, transcribe, and analyze online community radio broadcasts in Eastern DRC, a region where radio remains the primary information source due to low internet penetration. The core problem addressed is the rampant spread of misinformation, disinformation, and hate speech through radio channels, which traditional manual monitoring methods cannot effectively mitigate.

Existing solutions, such as manual transcription and feedback from listener groups, are slow, labor-intensive, and often inaccurate. Conventional speech-to-text tools struggle with the region's mixed languages, dialects, and high-noise environments, leading to significant transcription errors. These limitations delay the detection of harmful content and impede timely responses to crises, thereby exacerbating social tensions.

Our AI-based approach leverages advanced speech-to-text technology optimized for French and Swahili, coupled with generative AI models that analyze transcribed content in near real-time. This integration facilitates scalable monitoring across hundreds of radio stations, enabling swift identification and classification of harmful content. By automating these processes, Radiolytica not only accelerates the detection of dangerous narratives but also supports the rapid production and dissemination of factbased counter-content.

We follow an oN-the-shelf approach by using the latest models released by industry leaders, ensuring we stay up to date and deliver the best possible results. This strategy also helps us avoid expensive fine-tuning of models and keeps our focus on the essence of the project: detecting and countering harmful narratives in real time.

While the AI approach has significant benefits—scalability, speed, and efficiency—it also faces challenges. Achieving high transcription accuracy in a multilingual, noisy environment remains a technical hurdle. There is also a risk of AI-generated misclassifications or "hallucinations," which could undermine trust in the system. To mitigate these drawbacks, our solution incorporates a hybrid model that combines automated processing with human validation by local experts and journalists. This ensures that the insights generated are both contextually relevant and reliable, ultimately enhancing operational eNectiveness and fostering social cohesion in conflict-affected regions.

**Use Case Status:** The use case is part of a larger research project

### Partners

Fondation Hirondelle [1] has a network of more than 150 partner radio stations across the DRC, including around 60 in the east of the country. It also has a media outlet based in Kinshasa called Studio Hirondelle RDC.

**ETH Zurich**: Our well-established collaboration with the Development Economics Group (DEC) at ETH Zurich helps us address challenges in defining relevant analytical questions and ensuring academic rigor. Their research expertise and methodological support enhance project impact. Together, we apply for grants, including an ongoing academic grant for the proposed project, focused on development economics questions.

**ICRC**: Preliminary discussions have highlighted the relevance of our work in humanitarian contexts and we have a memorandum of understanding with ICRC. We aim to build on these initial exchanges. Their role is to provide iterative feedback during the development phase and serve as early adopters once we expand.

## 2.2 Benefits of the Use Case

Radiolytica directly tackles the destabilizing effects of misinformation, disinformation, and hate speech that undermine societal trust. By providing an automated, scalable solution for monitoring harmful content on community radio, the project enhances the integrity of information flows, supports accountable institutions, and fosters peace in conflict-affected regions.

Leveraging cutting-edge AI, advanced speech-to-text, and NLP technologies, Radiolytica modernizes media monitoring in low-resource settings. This innovative digital infrastructure not only addresses current communication challenges but also creates a resilient system capable of evolving with future technological advancements, contributing to sustainable industrial development.

Radiolytica is underpinned by a collaborative framework that unites global tech experts, academic institutions, and humanitarian organizations with local Congolese partners. By involving Congolese organizations, the project ensures that local context and expertise drive the development and application of AI, promoting inclusivity and capacity building within the region. This approach fosters mutual learning, strengthens local innovation ecosystems, and aligns with the vision of building strong, inclusive partnerships that drive sustainable development.

## 2.3 Future Work

Our future work will build on our successful pilot to transform Radiolytica into a fully scalable system for monitoring harmful content across Eastern DRC's community radio networks.

***Scaling and Refinement of AI Capabilities:***

We will expand our prototype to monitor up to 150 radio stations, enhancing automated transcription and NLP analysis. Our focus is on refining our pre-process pipeline for French and Swahili, reducing errors from high noise levels and overlapping audio. With additional resources, we also aim to explore incorporating Lingala, broadening our language coverage.

***Enhanced Human-AI Hybrid Validation:***

A core aspect is our hybrid model that pairs automated processing with continuous human validation. Local experts, journalists, and media partners will verify AI outputs, ensuring contextually accurate, reliable data and mitigating risks like AI hallucinations.

***Local Partnerships and Capacity Building:***

Leveraging collaborations with the Development Economics Group at ETH Zurich and ICRC, we will strengthen partnerships with Congolese organizations. This will enhance local expertise in AI model training and promote **capacity building in digital transformation, ensuring our solution is both inclusive and sustainable.**

***Exploration of Automated Streaming Solutions:***

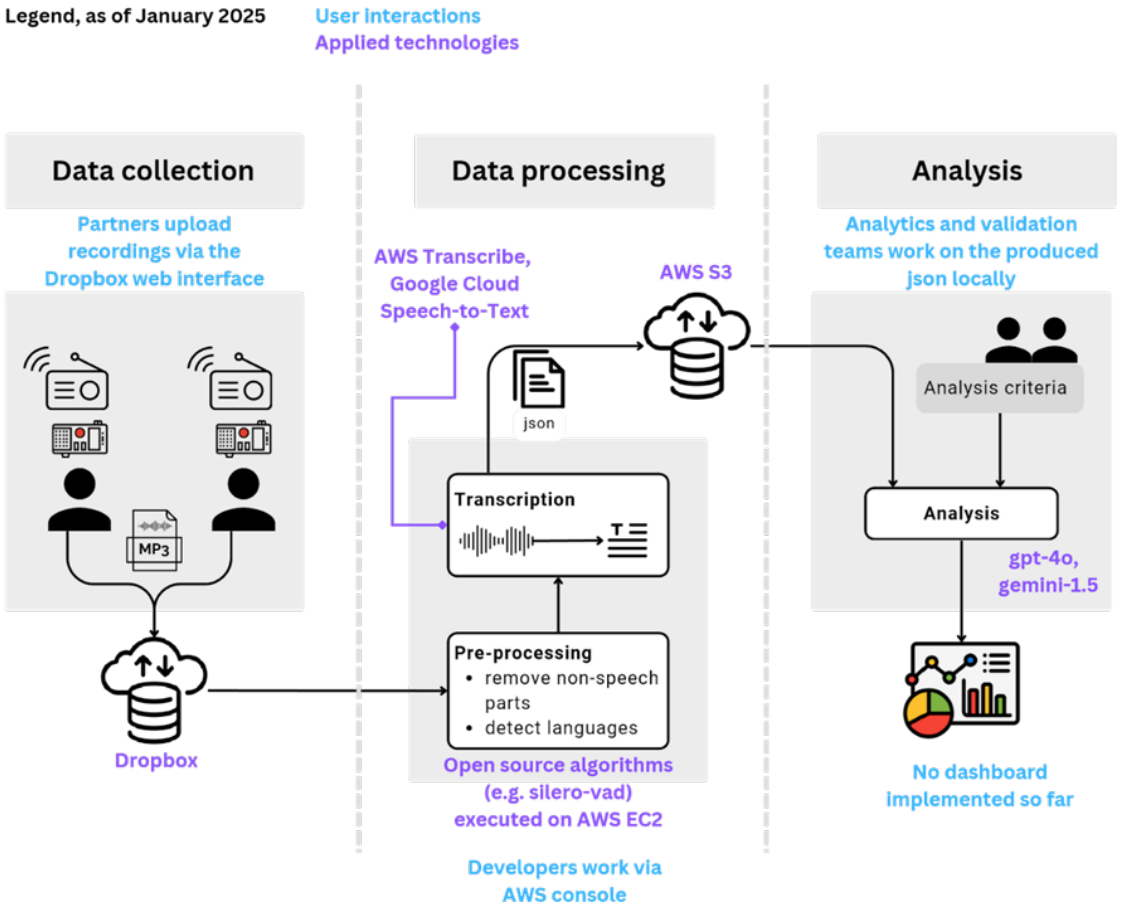
Further feasibility studies on automating data collection using Raspberry Pi devices will be conducted in regions with stable infrastructure, addressing trust issues noted in our initial analysis.

This roadmap will transform Radiolytica into a trusted, scalable system that mitigates harmful content and strengthens social cohesion.

## 3 Use Case Requirements

- **REQ-01**: It is critical to collect data from community radio stations. This requirement is met through Fondation Hirondelle's established network across Eastern DRC, which ensures consistent and reliable access to relevant audio content.
- **REQ-02**: It is critical to develop and maintain the data processing pipeline. It is essential to manage the integration of speech-to-text services, generative AI models, and the hybrid validation architecture.
- **REQ-03**: It is expected that the speech-to-text and large language model (LLM) tools support pay-as-you-go basis, making sustained funding critical for scalability and long-term deployment.
- **REQ-04**: It is critical that AI-generated outputs are validated by experts. This human-in-the-loop component ensures contextual accuracy and mitigates the risk of AI errors. Fondation Hirondelle's in-house staN and regional partners provide the necessary linguistic and cultural expertise.

## 4 Sequence Diagram




## 5 References

[1] Fondation Hirondelle https://www.hirondelle.org/en

[2] Interview with Gilles Magnin, programme manager at Fondation Hirondelle https://information.tv5monde.com/international/video/comment-lafrique-sempare-delia-2762297

# Use Case 6: Artificial Intelligence-Generated Content Moderation

浙江-法国数字媒体取证联合实验室
Sino-French Joint Laboratory for Digital Media Forensics of Zhejiang Province

杭州电子科技大学
HANGZHOU DIANZI UNIVERSITY

**Country**: China

**Organization**: Hangzhou Dianzi University

**Department/Division/Group**: School of Cyberspace, Sino-French Joint Laboratory for Digital Media Forensics of Zhejiang Province

**Contact Person(s):** Tong Qiao, tong.qiao@hdu.edu.cn,

## 1 Use Case Summary Table

| Item | Details |
|---|---|
| Category | AI Security; Content Moderation; Multimedia Forensics |
| Problem Addressed | Content moderation for AIGC, focusing on hallucinations, synthetic content, and deviations from established standards to enhance the security and reliability of AIGC content. |
| Technology Keywords | Unsupervised learning, data augmentation, multi-level alignment mechanism, multimodal large language models, generative models, AIGC moderation |
| Data Availability | Data is publicly available |
| Testbeds or Pilot Deployments | At the Inclusion The Global Multimedia Deepfake Detection Challenge, the detection accuracy reached approximately 95%.<br><br>Applied in the Public Security Bureaus in China to combat telecom fraud involving the use of artificial intelligence. |

## 2 Use Case Description

### 2.1 Description

This case aims to address issues such as hallucination in large language models, the generation of fake content, and AIGC that does not meet expected standards. For instance, the European Commission's AI Act (2024), the world's first comprehensive regulation on AI, requires AI technology providers to ensure that synthesized audio, video, text, and image content can be detected as AI-generated. The AI Act also prohibits the use of AI systems deemed to pose a clear threat to fundamental user rights. Relevant regulations include the European

Commission's GDPR (2016), China's Interim Measures for the Management of Generative Artificial Intelligence Services (2023), and the Global AI Governance Initiative (2023). This work focuses on moderating AIGC content to ensure compliance with regulations across different countries and regions, and covers the academic and industrial AIGC products such as GPT-4o, Imagen 4, Midjourney, Stable Diffusion, Veo 3, Gen-4, Pika, Sora, ACE-Step, Stable Audio, MuseNet, V2A, DeepSeek, Gemini, Claude, Llama 4.0, etc. The solution will leverage the advanced artificial intelligence to develop an App, platform, SDK and mini-program to verify the authenticity and compliance of AIGC content. Existing solutions primarily rely on manual review or rule-based keyword engines, but these methods are limited by the complexity of contextual understanding and the ability to process vast amounts of data. For instance,

- Content moderators manually review text, images, and videos to determine whether the content violates platform policies, such as those related to political sensitivity, pornography, violence, or misinformation. However, this approach is highly subjective and costly, making it difficult to scale effectively for large volumes of content.
- The regular expression matching method identifies patterns within input text by defining a set of rules or patterns using regex syntax. However, it has notable limitations, such as a lack of semantic understanding and poor readability and maintainability when dealing with complex regular expressions.
- The DFA (Deterministic Finite Automaton) sensitive word filtering algorithm efficiently identifies sensitive terms in a given text by matching them against a predefined keyword database. However, the algorithm has limitations, including a lack of semantic understanding and an inability to handle complex contextual relationships, which makes it vulnerable to evasion.

AI-based approaches offer several benefits:

(1) They can handle massive data volumes simultaneously, significantly improving detection efficiency.

(2) They capture contextual semantic relationships, enhancing detection accuracy.

(3) They integrate multimodal features, further improving detection accuracy.

However, AI-based methods also have drawbacks:

(1) They depend heavily on large datasets with reliable labels, making them vulnerable to data poisoning attacks and prone to poor generalization.

(2) They lack robustness against post-processing operations like compression and cropping, which are commonly used on social platforms.

(3) The features extracted by AI models often lack interpretability.

To address the first drawback, this work adopts unsupervised learning to reduce dependence on large amounts of labeled data and to resist data poisoning attacks. Additionally, by constructing a diverse AIGC database that covers various types, styles, and scenarios, the generalization capability of the model is significantly improved.

To tackle the second drawback, this work integrates multimodal features, including audio, visual, and textual modalities, and applies data augmentation to enhance the model's robustness against common post-processing operations such as compression and cropping.

To overcome the third drawback, this work builds an AIGC moderation platform and designs a multi-level alignment mechanism for safety standards, thereby enhancing the regulatory compliance of generated content. Moreover, by leveraging the powerful capabilities of

multimodal large language models, the platform not only detects content authenticity but also identifies manipulated regions, generates heatmap-based visualizations and provides textual explanations of the model's judgment criteria to improve interpretability and user trust.

## Proposal

Firstly, based on mainstream generative models such as Midjourney, GLIDE, and Stable Diffusion, this case will construct a diverse AIGC dataset. Firstly, this work leverages mainstream generative models to achieve comprehensive multimodal content coverage. For the text modality, models such as DeepSeek V3 and Llama 4.0 are used to generate diverse types of content, including news, commentary, and popular science articles. For the image modality, models like GPT-4o and Imagen 4 are employed to produce images with varying resolutions, subjects, and artistic styles. For the video modality, generative models such as Veo 3 and Gen-4 are utilized to synthesize videos across different scenes and artistic styles. For the audio modality, models like ACE-Step and Stable Audio are used to generate speech with multilingual, gender-diverse, and emotionally varied characteristics. Secondly, a wide range of prompts is designed to enhance the diversity and complexity of the generated content, covering various cultural contexts, neutral and sensitive topics, as well as high-risk and ethically challenging scenarios. Finally, each sample is accompanied by detailed metadata, including the source model, input prompt, content type, post-processing operations, and multimodal association identifiers, ensuring the dataset's utility for downstream moderation and forensic research. The dataset includes various content types, styles, and scenarios, providing rich training data to support subsequent moderation research. Next, this case adopts an unsupervised domain adaptation strategy, extracting and integrating multimodal features (text, image, audio, etc.) to study general multimodal forgery detection methods, enhancing both the accuracy and applicability of detection. Additionally, the case designs a multi-layered safety standard alignment mechanism to balance global commonality with regional characteristics. The first layer involves globally universal standard moderation, while the second layer focuses on region-specific customized standard moderation. Through this hierarchical mechanism, it ensures both global compliance of AIGC content and flexibility in adapting to the cultural and social needs of different regions. Finally, based on the forgery detection and content moderation models, an AIGC moderation platform will be developed. This platform will collaborate with social media platforms to promote its demonstration application and industrialization. The platform will continuously update its data and iteratively improve model performance based on test results. Moreover, it will support customized moderation for different scenarios. Furthermore, the case will apply for industry standards to increase the adoption and influence of AIGC moderation.

- Content Moderation: Social media platforms can utilize the proposed AIGC moderation platform to detect and identify misinformation, implement appropriate countermeasures to limit its spread, and protect users from deception and harm. This helps to ensure the authenticity and credibility of disseminated AIGC content, fostering a safe and trustworthy online environment for users.
- Figure Protection: The proposed AIGC moderation platform is capable of detecting forgery videos or audio recordings of political leaders, as well as identifying maliciously altered images, fabricated news, and misleading statements. By preventing the spread of such content during major political events such as elections, the platform helps to safeguard public opinion, prevent misinformation, and protect national reputation and political stability.
- Public Security: Regulatory authorities can leverage the proposed AIGC moderation platform to detect security risks and malicious activities caused by synthetic content. This

includes identifying and preventing the spread of disinformation, combating telecom fraud and false advertising, promptly detecting the forgery of public figures' speech or likeness, and taking timely countermeasures. Such capabilities are crucial for safeguarding social stability, national security, and public order.

- Military Security: The proposed AIGC moderation platform is capable of effectively detecting falsified military intelligence, such as fabricated battlefield images, manipulated satellite photos, and fake videos or announcements of military deployments. By identifying such synthetic content, the platform helps to prevent hostile actors from misleading command decisions, disrupting troop operations, or inciting public panic, thereby safeguarding national military secrecy and ensuring the stability of defence strategies.
- Financial Security: The proposed AIGC moderation platform can detect fraudulent activities involving the use of forged facial images or synthetic speech to bypass facial or voiceprint identity verification systems, thereby preventing telecom fraud, unauthorized fund transfers, and related schemes. Additionally, the platform is capable of identifying falsified financial announcements, investment advice, and other misleading content, helping to prevent public misinformation and disruptions to financial order.
- Academic Integrity: The proposed AIGC moderation platform is capable of detecting the misuse of generative AI technologies in academic writing and assignment submissions. It enables timely identification and prevention of academic misconduct, thereby ensuring fairness in educational assessment and upholding academic standards and educational quality.
- Risk Control: The proposed AIGC moderation platform can detect malicious activities on social media platforms that exploit generative AI technologies, such as the mass creation of fake accounts, fabrication of false personas, dissemination of misinformation, manipulation of trending topics, and engagement in fraudulent schemes. The platform enables timely identification of sources of public opinion risks, curbs the spread of rumours, and combats information manipulation.
- Community Security: Smart community platforms can leverage the proposed AIGC moderation platform to identify and curb the spread of false information, maintain community order, and reduce fraudulent activities such as deceptive advertisements and scams. Additionally, it helps to protect users' privacy and personal image, fostering a trustworthy and secure community environment.

## 2.2 Benefits of the use case

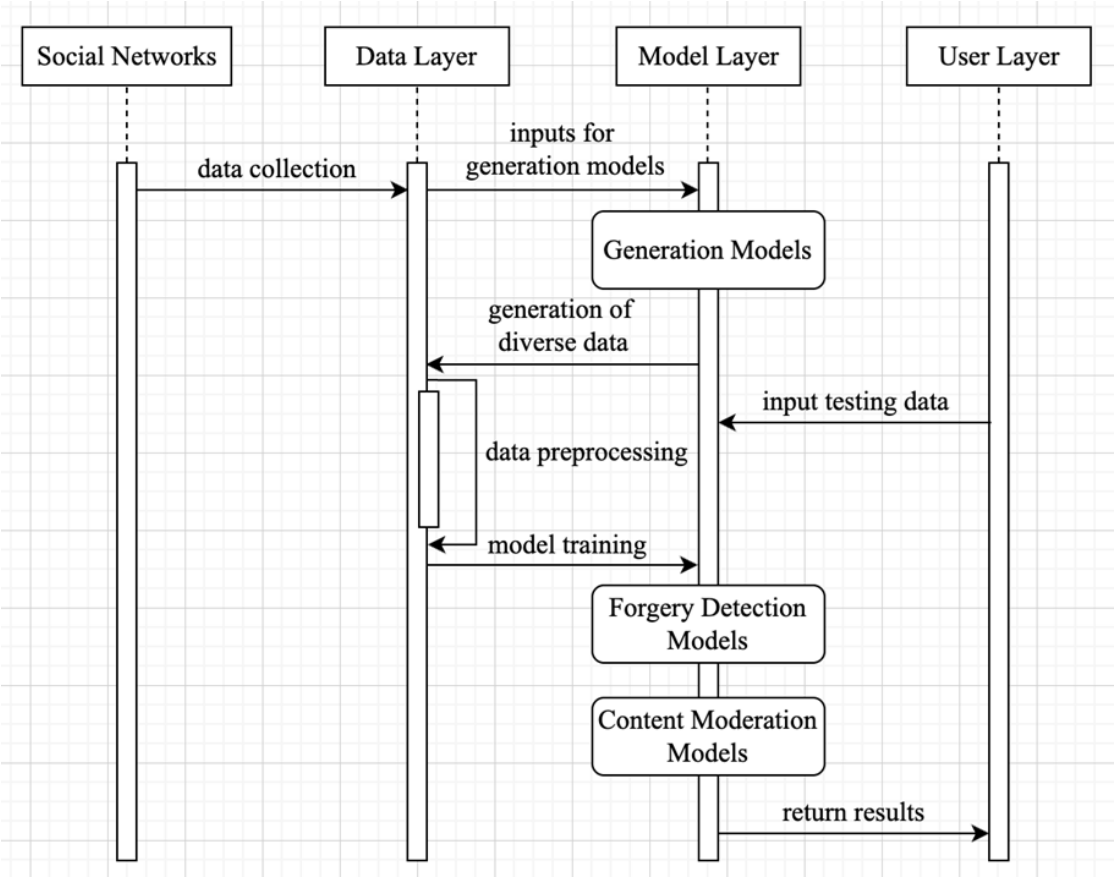
This case leverages artificial intelligence technology to verify the authenticity and ensure the compliance of AIGC content, enhancing the credibility of digital content, reducing its sensitivity, and ensuring it aligns with global safety standards, thereby promoting a healthy and safe information ecosystem. In a globalized digital environment, the spread of fake or misleading content can impact public perception, exacerbate social divisions, and challenge the credibility and governance capabilities of institutions. This case can efficiently and accurately identify fake information and generated content that does not meet safety standards, reducing the spread of disinformation and mitigating its associated risks and harms. The technology can enhance the efficiency and consistency of content governance within organizations, reduce the costs and subjective biases of manual audits, and contribute to the creation of a safe global digital ecosystem.

## 2.3 Future work

- Data collection
- Proof of concept development

- Model development
- Create new variations/extensions to the same use case
- Standards development related to the use case
- Set up reference tools, notebooks, and a simulation
    - Completed in the first half of 2025: A diverse AIGC dataset is constructed based on mainstream generative models to provide comprehensive data support for subsequent content moderation tasks. An unsupervised domain adaptation strategy is employed to extract and integrate multimodal features for the development of general multimodal forgery detection techniques. In parallel, a multi-level safety standards alignment mechanism is designed to balance global commonalities with regional specificities. This hierarchical approach ensures global compliance of AIGC content while accommodating the cultural and societal norms of different regions. Finally, forgery detection models and content moderation models are trained.
    - Completed in the second half of 2025: An AIGC moderation platform is developed based on the proposed forgery detection and content moderation models. Leveraging this platform, pilot applications and industrial-scale deployments are carried out in collaboration with local public security authorities and technology enterprises in China. The platform supports continuous data updates and iterative model improvements based on real-world testing outcomes. Efforts are also made to propose industry standards, promoting the widespread adoption and influence of AIGC content moderation technologies. In addition, open-source tools, guidance documents, and simulation settings are provided to support broader implementation and research.

## 4 Sequence Diagram



## 5 References

[1] Qiao T, Xie S, Chen Y, et al. Fully unsupervised deepfake video detection via enhanced contrastive learning[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2024, 46(7): 4654-4668.

[2] Lyu S. DeepFake the menace: mitigating the negative impacts of AI-generated content[J]. Organizational Cybersecurity Journal: Practice, Process and People, 2024, 4(1):1-18 (ahead-of-print).

[3] Zhu M, Chen H, Yan Q, et al. Genimage: A million-scale benchmark for detecting ai-generated image[J]. Advances in Neural Information Processing Systems, 2023, 36: 77771-77782.

# Use Case 7: AI model to summarize, rate, and flag potentially harmful clauses found in terms of service or privacy policy agreements for digital services

**Country**: Zambia

**Contact Person(s)**:

Nchimunya Zoe Kabalo, nzkabalo@gmail.com

## 1 Use Case Summary Table

| Item | Details |
|---|---|
| Category | Digital consumer services, legal contracts |
| Problem Addressed | Many consumers agree to, but do not read, the ToS of the various services they use due their complexity and length, leading to them agreeing to terms who's implications they may not fully understand, and that could potentially violate their user and privacy rights. |
| Key Aspects of Solution | AI model to summarize and rate ToS and privacy policies<br>Mobile application for convenient access and a more personalized use |
| Technology Keywords | Large Language Model (LLM)<br>Terms of Service (ToS)<br>Privacy policy |
| Data Availability | Public |
| Metadata (Type of Data) | Text based labeled data containing ToS and Privacy Policy documents, their ratings, their summaries, their clauses and the category of those clauses (good, bad, critical) |
| Model Training and Fine-Tuning | Natural Language Processing (NLP) is used to read and provide a summary the ToS of a service. Machine learning is used to enable the AI model to identify and differentiate between fair and unfair clauses on the basis of how well consumers' rights and their privacy is protected. |
| Testbeds or Pilot Deployments | N/A |
| Code repositories | N/A |

## 2 Use Case Description

### 2.1 Description

Though containing critical rules and conditions that can help to protect the rights of both the service provider and the user, the Terms of Service(ToS) agreements of most services and platforms these days have a reputation for being notoriously long, complex, and filled with legal jargon, making it difficult for the average user to understand their implications (Obar and

Oeldorf-Hirsch, 2016). As a result, many users blindly agree to these terms without reading or comprehending them (Milne and Culnan, 2004), potentially compromising their privacy or unknowingly agreeing to unfavorable clauses. This creates an imbalance of power between users and service providers, allowing companies to draft ToS agreements that heavily favor their interests while limiting user rights and imposing strict liabilities on them (Marotta-Wurgler, 2010). With the increasing use of digital services, this problem affects a vast number of users worldwide.

Artificial intelligence (AI) offers a solution at combating the issues described above. Due to its automation and speed, AI has the ability to process and analyze large volumes of documents faster and more accurately than humans can. Also, AI systems have the ability to improve over time by learning from user interactions and past performance. These features make AI a useful tool in helping everyday users to understand and analyze terms of service in ways that have previously proved difficult. Additionally, Machine Learning (ML) can be used to help the model recognize patterns in legal documents, identifying problematic clauses or inconsistencies across multiple contracts (Surden, 2014; Katz et al., 2014). Over time, it can become more accurate at spotting unfair terms or data privacy concerns.

While AI tools for legal documents, such as LawGeex, exist, there remains a lack of AI-driven tools specifically focused on summarizing terms of service for everyday users. These tools, while powerful, are often tailored for businesses or legal professionals, leaving a gap in the market for tools that cater to individual consumers. Current literature (Goodman and Flaxman, 2017) underscores that while AI is improving the legal domain, consumer protection remains underdeveloped in AI research. The need for AI tools that simplify legal documents for the general public, and everyday users, who often lack the time or expertise to analyze contracts, remains largely unmet. AI systems that can offer real-time analysis of terms of service, flagging risky clauses or comparing terms across services, could play a critical role in empowering users.

The proposed use case aims at utilizing these properties and technologies of AI in order to create an application with an AI model that can summarize, rate and flag potentially unfair clauses contained in the ToS agreements of various digital services in order to provide better understanding and transparency, empowering users to make more informed decisions and better protect their digital rights and privacy. The AI model will be a pre trained legal LLM fine tuned on a comprehensive ToS dataset to perform the following major functions:

- summarize ToS and privacy policy documents
- rate ToS and privacy policy documents based on how well user rights are protected
- identify and flag potentially problematic or unfair clauses

**Use Case Status**: The use case is part of a final year research project

## 2.2 Benefits of the Use Case

- The proposed use case can address key gaps in fields such as consumer protection, ToS transparency, and the application of AI in legal domains, which could help to further advance the research in and development of these fields.
- The proposed use case encourages ethical consumption by educating users about their right and implications of ToS and privacy policy agreements as well as encouraging them to make informed choices about the digital services they use. It can also foster responsible business practices by encouraging producers of these digital services to review their terms of service and how they handle user data.

- The proposed use case promotes transparency by helping users understand the legal terms they agree to, ensuring that companies are held accountable for unfair or exploitative clauses. By flagging problematic clauses and providing clear summaries, users are empowered to make informed decisions and seek legal recourse if necessary. The proposed use case also aligns with the goal of building strong institutions by advocating for fair and ethical business practices.

## 2.3 Future Work

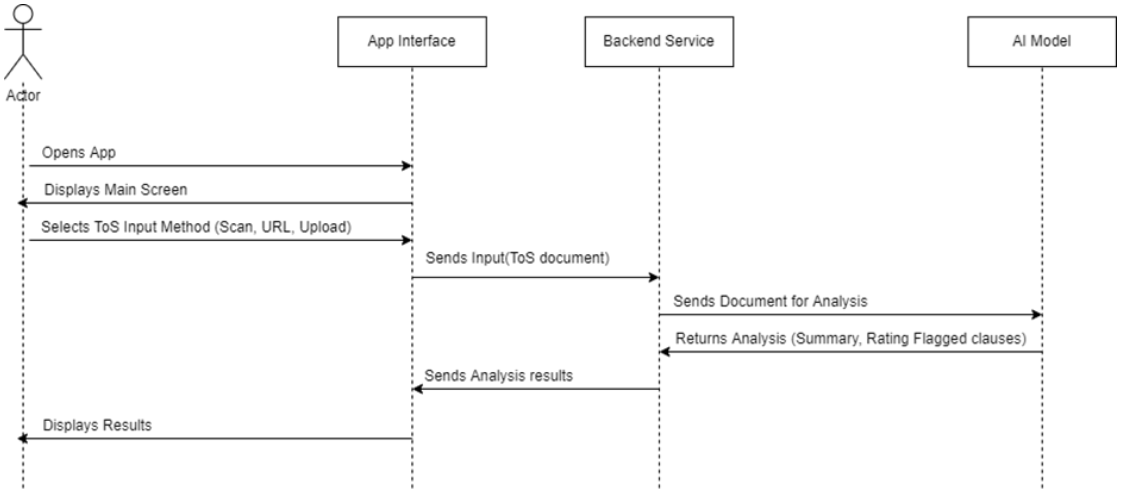
Given more resources, more work can be done in terms of making a much more comprehensive data set for fine tuning the AI model. More work will be put in to collecting more data for the data set. This includes ToS and privacy policy documents of more digital services, and acquiring professional help in summarizing such documents. The more comprehensive data set will then be used to further fine tune the AI model in order to make it more accurate and provide more trustworthy summaries to users.

Additionally, more work can be done in making the resulting mobile application even more helpful and educational to users. Features such as comparing similar applications so as to recommend the one with more favourable ToS, providing users with more educative information on clauses, and helping users know what to look out for before agreeing to terms, will be added in order to achieve this.

# 3 Use Case Requirements

- **REQ-01**: Create a comprehensive data set with ToS and privacy policy documents of various digital services, their ratings, their summary, and their good and bad clauses. This can be reviewed and standardized with the help of experts.
- **REQ-02**: Fine tune a pre trained legal LLM in NLP and ML using the created data set to be able to read and summarise ToS documents and identify clauses
- **REQ-03**: Validate the output of the AI model
- **REQ-04**: Create a mobile application into which the model can be integrated and allow users to interact with it

# 4 Sequence Diagram



## 5 References

[1] Obar, J. A., & Oeldorf-Hirsch, A. (2016). The biggest lie on the internet: Ignoring the privacy policies and terms of service policies of social networking services.

[2] Milne, G. R., & Culnan, M. J. (2004). Strategies for reducing online privacy risks: Why consumers read (or don't read) online privacy notices.

# Use Case 8: Uncovering Delays: Leveraging Data Science and AI to Streamline Refugee Status Determination Procedures

**Organization:** United Nations High Commissioner for Refugees (UNHCR)

**Country:** Switzerland

**Contact Person(s)**:

Adriana Rojas, rojasa@unhcr.org

Moises Maldonado maldonam@unhcr.org

Peter Grady, grady@unhcr.org

## 1 Use Case Summary Table

| Item | Details |
|---|---|
| Category | Productivity |
| Problem Addressed | UNHCR conducts refugee status determination (RSD) procedures under its mandate in approximately 50 countries, when there is an absence of fair and efficient asylum procedures.<br>Unprecedented Levels of forced displacement: In 2024, UNHCR received 484,000 new individual asylum applications, raising to 1,173,200 the number of Asylum-Seekers registered with UNHCR without a final decision (UNHCR, 2025).<br>Lengthy Processing Times: the RSD process, including interviewing, assessment and reviewing stages, can vary in length from one day to several months.<br>Wide Variations: Time taken for final decision on case resolution and notification varies widely due to countries operational context, processing strategies, methods, profiles, caseloads, and resource availability.<br>Need for Analysis: Stark variations in processing times highlight the need for detailed data analysis to identify inefficiencies and streamline asylum procedures, adapting it to the particular countries and needs. |

(continued)

| Item | Details |
|---|---|
| Key Aspects of Solution | Data mining: extracting anonymised data from UNHCR's case management system.<br>Exploratory data analysis:<br>Process mining – map the process from the moment of the asylum application until a decision is made to identify patterns.<br>Big data analysis – analyse log data to understand bottlenecks (demographics and operational).<br>Machine learning (ML) and artificial intelligence (AI) to identify trends: train the model using ML and AI algorithms to identify bottlenecks, identify root-causes of delays, and predict the length of the RSD process.<br>Prescriptive analysis for efficiency gains: provide recommendations to avoid delays, gaining efficiency in the RSD process under UNHCR's mandate. |
| Technology Keywords | Data mining, process mining, machine learning, artificial intelligence, linear regression modelling, predictive analytics. |
| Data Availability | The data source is private. Anonymised and aggregated data is extracted from UNHCR's case management system following privacy-by-design and data-minimisation principles in compliance with UNHCR's data protection and privacy framework (UNHCR, 2022). Data is stored in a secure Microsoft Fabric environment (subject to a framework agreement with UNHCR). |
| Metadata (Type of Data) | Text: Relational database, structured SQL tables, process logs, categorical data, text data. |

(continued)

| Item | Details |
|---|---|
| Model Training and Fine-Tuning | Several AI and ML techniques are employed to identify key factors contributing to delays in the RSD process. These include:<br>Linear Regression: A statistical ML technique used to quantify the relationship between processing time and various predictors such as case characteristics, resource allocation, or workflow stages. It helps in identifying which variables are significantly associated with longer delays.<br>Regression Trees and Decision Trees: These tree-based models provide a visual and interpretable way to understand how different combinations of variables (e.g., country of origin, type of procedure, staff availability) contribute to variations in processing time.<br>Random Forests: An ensemble learning method that builds multiple decision trees and aggregates their outputs for improved prediction accuracy. Random forests help in ranking the most important predictors of delays and capturing complex patterns that may not be evident in simpler models.<br>Model Fine-Tuning plays a critical role in ensuring that these ML models are both accurate and generalisable. Techniques such as cross-validation, hyperparameter optimisation, and regularisation are used to adjust the models for better performance. This process helps avoid overfitting, improves predictive accuracy, and ensures that the insights derived from the models are robust and actionable.<br>In addition to these ML methods, AI-driven process mining is used to analyse event logs and uncover the actual workflows within the RSD process. This technique provides insights into bottlenecks, inefficiencies, and deviations from standard operating procedures, offering a dynamic, data-driven view of how cases move through the system.<br>To avoid the risk of exposing personally identifiable information through reverse engineering of our ML model, particularly through the data used to trained it, data will be encrypted and some variables eliminated from the model. This will provide a minimum dataset (data minimisation) with enhanced privacy and data protection. |
| Testbeds or Pilot Deployments | Pilot deployments are expected to take place in the next stage of the project, targeting specific locations in which RSD is conducted across a range of geographical areas. |
| Code repositories | Not available |

# 2 Use Case Description

## 2.1 Description

RSD is the process through which governments or UNHCR decide if a person seeking international protection qualifies as a refugee under international, regional, or national law. This process is crucial for ensuring that refugees can exercise their rights under international law. Each year, UNHCR carries out RSD procedures under its mandate in around 50 countries where fair and efficient asylum systems are not in place.

After the registration and recording of the applicant's biodata, the RSD process under UNHCR's mandate includes three key stages:

**interviewing stage**, where information is collected from the Applicant on their refugee claim,

**assessment stage**, where eligibility officers make a legal analysis to determine whether the applicant meets the refugee definition under the 1951 Convention and its protocol, and

**reviewing stage**, where the legal analysis is reviewed by a supervisor and either endorsed or returned for re-interview or redraft of the legal analysis.

Despite efforts to ensure a timely issuance of decisions through the implementation of varying processing modalities and strategies and the application of procedural standards for mandate RSD (UNHCR, 2015; UNHCR, 2020), the duration of the process can range from a day to many months.

Lengthy RSD processing times pose significant challenges, particularly for Asylum-Seekers awaiting decisions on their applications. The duration of these procedures can significantly impact the well-being and safety of Asylum-Seekers (Phillimore et al., 2021). Uncertain refugee and resulting legal status can hinder access to essential services, expose individuals to human rights breaches, and delay access to long-term solutions.

Effective and timely decisions on refugee claims enable recognised refugees to find protection and pursue socio-economic inclusion in their country of asylum. Additionally, prompt decision-making supports the safe, dignified, and rights-based return of individuals who do not qualify for international protection, reducing the misuse of asylum procedures and ensuring a better use of scarce resources.

The high number of asylum applications made with UNHCR coupled with dwindling resource, has led UNHCR to turn to data science to make an in-depth analysis of its RSD process and leverage new technologies, including process mining and ML techniques, to identify both bottlenecks and efficiencies in the RSD process.

With advanced data modelling and AI-powered tools, UNHCR aims to uncover root causes of delays at each of the stages in the RSD process (interviewing, assessment and reviewing) under its mandate, and devise ways to tackle such delays or prevent them altogether. This analysis will also allow the identification of efficiencies in the RSD procedure, which could be replicated and streamlined across UNHCR RSD procedures.

By combining advanced ML techniques with AI-driven process mining, the technological approach identifies causes of delays in UNHCR's RSD process. This integrated approach leverages predictive models - such as linear regression, decision trees, and random forests—to identify key drivers of processing delays, while also using process mining to visualise real-time workflows and detect inefficiencies or bottlenecks in the system. The models are fine-tuned through cross-validation and hyperparameter optimisation to enhance accuracy and reliability. By uniting predictive analytics with dynamic process mapping, this solution offers a powerful, data-driven framework for improving operational efficiency and guiding evidence-based decision-making in refugee case management. Process mining can be viewed as a post-facto analytical approach that examines event log data to uncover actual process flows, bottlenecks, and variations within a system. In the context of the RSD process, it enables the identification of patterns in the duration and progression of individual cases through various stages. By reconstructing the real pathways followed, rather than relying solely on formal procedures or assumptions, process mining provides valuable insights into how the RSD process operates in practice. This retrospective analysis helps uncover systemic delays, inefficiencies, or deviations,

and supports a deeper understanding of the correlating elements—such as case characteristics, institutional workload, or procedural factors—that influence processing times. Ultimately, the goal is to use this knowledge to inform targeted improvements and policy decisions that enhance the efficiency and fairness of the RSD process.

Once concluded, the AI-powered analysis combining the process mining findings and ML modelling will be presented through a business intelligence report in the form of a Dashboard or WebApp designed for a non-technical audience. The business intelligence report will provide insights to UNHCR RSD managers on the possible root causes of delays and devise strategies to tackle and prevent them. Predictive modelling will give managers the tools to devise processing strategies to streamline RSD procedures enhance the efficiency of the RSD process and maximise the use of resources.

The projects will provide the UNHCR RSD Manager with the necessary information to develop data-driven RSD processing strategies and implement necessary changes, including a re-distribution of resources, to tackle and prevent bottlenecks in the future.

This project will contribute to the implementation of RSD procedures that are fairer, more efficient, flexible and adaptable. It will enable refugees to more swiftly access safety, solutions and live-saving services and contributing to the integrity of UNHCR's asylum system.

**Use Case Status**: In development.

**Partner**: N/A

## 2.2 Benefits of the Use Case

This project leverages advanced technologies to identify and resolve bottlenecks in UNHCR's RSD process, improving efficiency and delivering tangible benefits across poverty reduction, equity, and institutional strengthening.

**Reducing Poverty**

By streamlining the RSD process, the project enables faster access to legal status, which is a prerequisite for many essential services and long-term solutions—including voluntary repatriation, local integration, or third-country resettlement. Accelerated legal recognition helps individuals secure employment and rebuild their lives, reducing poverty among displaced populations. In turn, this contributes to economic growth and lowers the financial burden on host communities.

**Promoting Equity and Inclusion**

Timely and fair RSD decisions ensure that individuals of all backgrounds have equal access to international protection mechanisms. Legal recognition opens doors to critical public services such as education, healthcare, and employment. This fosters greater social and economic inclusion and helps reduce disparities between refugees and host populations.

**Strengthening Institutions and Governance**

The use of advanced tools like AI and process mining enhances the capacity of institutions to manage asylum procedures with greater speed, consistency, and accountability. The initiative not only supports UNHCR's operational mandate but also offers a replicable model for

national asylum systems and other processes serving displaced populations. By modernizing administrative procedures, the project helps build more resilient, transparent, and effective public institutions that safeguard human rights and international protection standards.

## 2.3 Future Work

The project is ongoing at the modelling stage. It is focused on developing and validating algorithms and training the data model.

In the next stage, if provided with additional resources, pilot countries will be identified to analyse the data in depth, validate the algorithms and consider other factors of delays not accounted for through data analysis.

A business intelligence report will be designed to combine the data generated from the process mining and AI model.

A dashboard/web app will be developed to present the results to a non-technical audience (decision makers) to address specific backlogs and provide policy recommendations.

Additional resources are required to implement the planned future work, including the development of a tool to automatise the end-to-end data extraction, data clean-up, data modelling and AI integration to generate the business intelligence report, and present the information in a clear and concise manner to a non-technical audience, particularly for the selected pilot countries aiming to potentially scale to the 50 country operations.

The project may be implemented across the organisation to benefit the approximately 50 country operations where UNHCR conducts RSD under its mandate. The integration of additional countries will enhance the capability of the tool identify efficiencies and propose new approaches across countries.

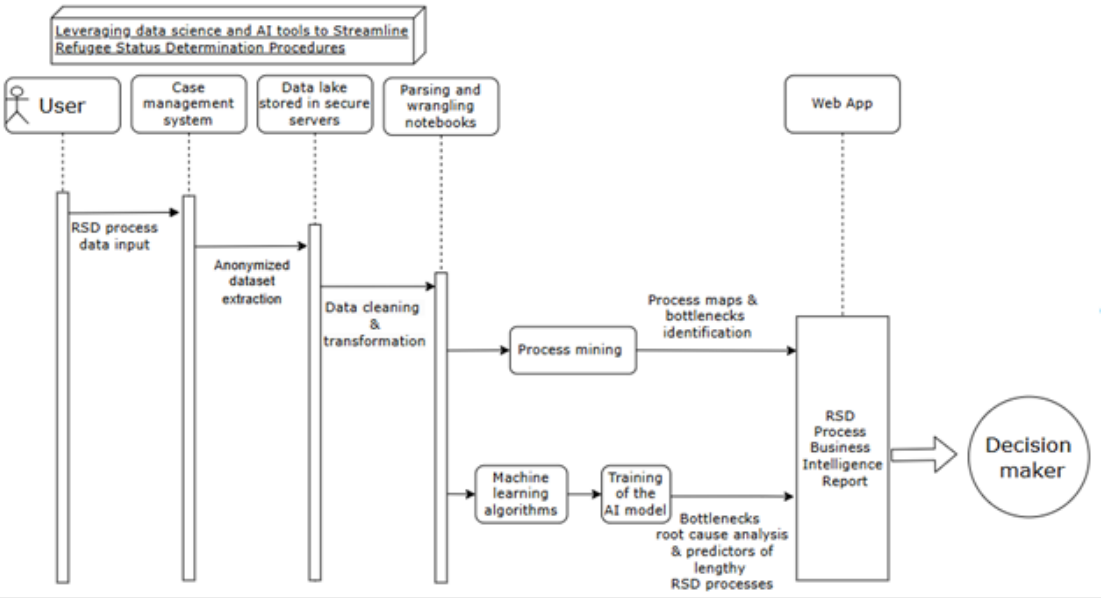
The research and model developed may also be adapted to RSD case management systems used by States in their national asylum procedures. With a backlog of over 8-million applications by mid-2024 and a global average processing time of approximately 400 days, this case study has great potential to create a global impact.

# 3 Use Case Requirements

- **REQ-01**: Case Management System with Privacy-Conscious Data Recording It is critical that a secure case management system is implemented to accurately capture RSD processing data. The system must include robust authentication protocols and adhere strictly to data privacy regulations, ensuring the confidentiality of sensitive individual information throughout the case lifecycle.
- **REQ-02**: Secure Internal Servers for Anonymized Data Storage It is critical that internal, access-controlled servers are available to store anonymized and aggregated data extracted from the case management system. This infrastructure is essential to protect data integrity and prevent unauthorized access, in alignment with data protection best practices.
- **REQ-03**: Scalable Big Data Processing Architecture It is expected that a high-capacity, scalable big data architecture is in place to handle large volumes of RSD-related data. This infrastructure must support real-time processing, efficient querying, and model training at scale, enabling timely insights and system responsiveness.

- **REQ-04**: Data Cleaning and Preprocessing Mechanisms It is expected that thorough data clean-up processes are conducted to ensure the relevance, consistency, and accuracy of extracted data. This step is vital to prevent model drift, bias, or misinterpretation during AI analysis and inference generation.
- **REQ-05**: AI Model Training and Fine-Tuning It is critical that dedicated resources and infrastructure are allocated for training and fine-tuning machine learning models using relevant, curated datasets. This step is essential to ensure that AI components are context-aware, reliable, and aligned with the specific decision-making needs of the RSD process.
- **REQ-06**: Web-Based Application with Interactive Dashboard It is of added value that a user-friendly web application is developed to display AI-generated inferences through a dashboard. The interface should include interpretation tools and visual analytics to support case officers, decision-makers, and institutional stakeholders in understanding and acting on model outputs with transparency and confidence.

## 4 Sequence Diagram




## 5 References

[1] UN High Commissioner for Refugees (UNHCR), Policy on the Protection of Personal Data of Persons of Concern to UNHCR, May 2015, https://www.refworld.org/policy/strategy/unhcr/2015/en/120873

[2] UN High Commissioner for Refugees (UNHCR), Procedural Standards for Refugee Status Determination Under UNHCR's Mandate, 26 August 2020, https://www.refworld.org/policy/legalguidance/unhcr/2020/en/123306

[3] UN High Commissioner for Refugees (UNHCR), General Policy on Personal Data Protection and Privacy, 20 December 2022, https://www.refworld.org/policy/strategy/unhcr/2022/en/124207

[4] UN High Commissioner for Refugees (UNHCR), Global Trends report 2024, 12 June 2025, https://www.unhcr.org/global-trends

[5] Ali, M. I. (2021). Why Am I Waiting? Analyzing Waiting Times in Business Processes from Event Logs (Extended Abstract). ICPM Doctoral Consortium.

[6] Phillimore, J., & Cheung, S. Y. (2021). The violence of uncertainty: Empirical evidence on how asylum waiting time undermines refugee health. Social Science & Medicine, 282, 114154. https://doi.org/10.1016/j.socscimed.2021.114154

[7] Arvidsson, M., & Noll, G. (2023). Decision Making in Asylum Law and Machine Learning: Autoethnographic Lessons Learned on Data Wrangling and Human Discretion. Nordic Journal of International Law, 92(1), 56–92, https://doi.org/10.1163/15718107-bja10057

[8] Chen & Eagle (2017), Can Machine Learning help predict the outcome of asylum adjudication? , ICAIL'17.

[9] Katsikouli et al. (2022), Machine Learning and Asylum Adjudications: From analysis of variations to outcome predictions, IEEE Access, Vol 10.

[10] Kinchin, N. (2021). Technology, Displaced? The Risks and Potential of Artificial Intelligence for Fair, Effective, and Efficient Refugee Status Determination. Law in Context. A Socio-Legal Journal, 37(3).

[11] Lashkevich, K., Milani, F., Chapela-Campa, D., Suvorau, I., & Dumas, M. (2024). Unveiling the causes of waiting time in business processes from event logs. Information Systems, 126, 102434, https://doi.org/10.1016/j.is.2024.102434

[12] Lashkevich et al. (2022), Why am I waiting? Data-driven analysis of waiting times in business processes, arXiv:2212.01392

[13] Memon et al. (2024), Artificial Intelligence (AI) in the asylum system, Medicine, Science and the Law 2024, Vol 64(2) 87-90.

[14] Ozkul (2023), Automating Immigration and Asylum: The uses of new technologies in migration and asylum governance in Europe, Oxford University: Refugee Studies Center.

[15] Piccolo et al. (2023), On predicting and explaining asylum adjudication, Mobile Working Paper Series, no.5.

[16] Ramji-Nogales et al. (2007), Refugee roulette: disparities in asylum adjudication, Stanford Law Review, Vol.60, Issue 2.

# Use Case 9: Developing an AI-powered translation app to improve workplace safety at construction and shipbuilding worksites

**Organization**: Crowdworks AI

**Country:** South Korea

**Contact Person(s)**:

Sue Kim, hs.kim@crowdworks.ai

Jung ho KIM, kimjuh@crowdworks.ai

Jeong Yeon KIM, waymaker@crowdworks.ai

## 1 Use Case Summary Table

| Item | Details |
|---|---|
| Category | Productivity |
| Problem Addressed | South Korea's construction industry employs over 100,000 foreign workers, a number that continues to grow. However, existing translation and interpretation tools are often inaccurate and fail to recognize the slang and terminologies commonly used on construction sites—leading to communication breakdowns and increasing the risk of serious accidents. |
| Key Aspects of Solution | The solution enables seamless real-time communication between workers at construction and shipbuilding worksites. Unlike the translation services currently available at the market, Crowdworks AI's solution accurate translates industry jargon, thereby preventing safety hazards. The solution utilizes below technologies:<br>• Multilingual corpus collection for construction and shipbuilding domain-specific translation model training<br>• Fine-tuning Gemma-2-9B-IT model [3]<br>• Applying PEFT via LoRA for efficient model training<br>• In-context learning and prompt engineering |
| Technology Keywords | LLM, Building Domain-Specific Datasets |
| Data Availability | Public (Glossary of National Construction Standards [1] by the Ministry of Land, Infrastructure and Transport (MOLIT)) + Private data from Korean construction sites |
| Metadata (Type of Data) | TEXT |
| Model Training and Fine-Tuning | Fine-tuning, Prompt-Engineering |

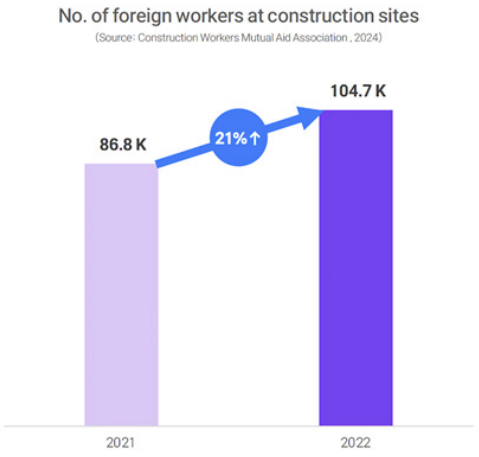
(continued)

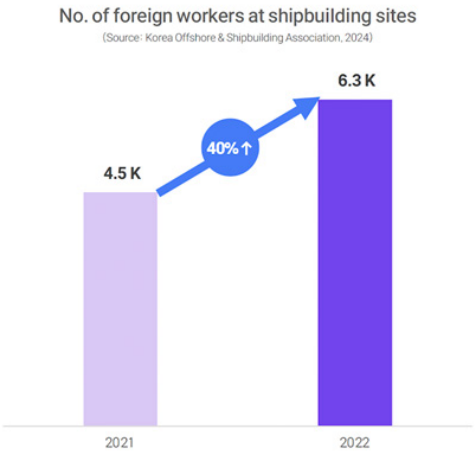
| Item | Details |
|---|---|
| Testbeds or Pilot Deployments | Data Availability:<br>• Public: Glossary of National Construction Standards by MOLIT [1]<br>• Private: Worker interviews and internal field data |
| Code repositories | Unavailable |

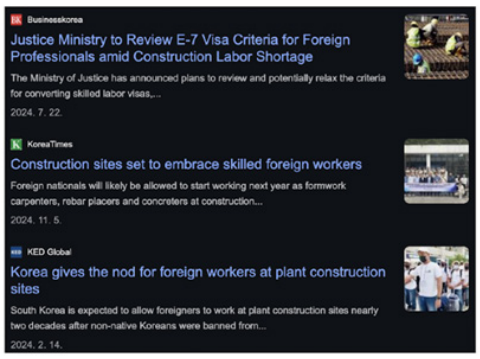
# 2 Use Case Description

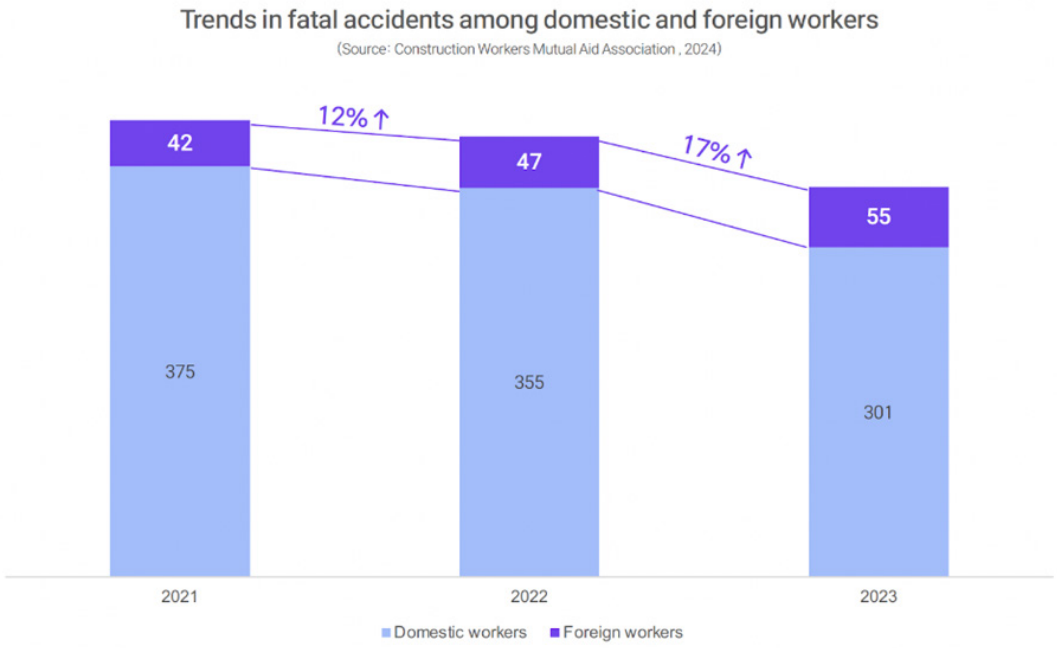
## 2.1 Description

As of 2022, over 100,000 foreign workers were employed in South Korea's construction and shipbuilding industry, a number that continues to grow.








To address this issue, a leading South Korean construction company S, implemented safety measures such as translating safety manuals and guidelines, placing safety posters in multiple languages at worksites, and hiring on-site interpreters. While these measures helped somewhat, they failed to ensure seamless real-time communication, and Company S concluded that to prevent safety risks, a real-time translation app is needed.

However, translation apps such as Google translate and Papago (a Korean-developed app widely used in South Korea's construction industry) failed to accurately translate industry-specific terms, leading to miscommunication and safety risks. To resolve this problem, Company Spartnered with Crowdworks AI to develop "Safety Talk," an AI-powered translation app specialized for the construction and shipbuilding industries. Safety Talk supports 22 languages, including English, Chinese, Spanish, Russian, and Arabic, ensuring seamless communication among multinational workers. The translator model addresses language variation and dialects through contextual fine-tuning and in-context learning. The model uses 100,000+ construction and shipbuilding domain-specific terms, designed and built to meet the precise needs of Safety Talk. Unlike the translation services currently available on the market, Crowdworks AI's solution accurately translates industry jargon, thereby preventing safety hazards.

**1 Building Doman-specific Dataset**

To mitigate confusion arising from lexical ambiguity, a domain-specific dataset tailored to the construction industry was developed

| Korean | Definition_Korean | Eng. | Definition_English |
|---|---|---|---|
| 개판 | 서까래나 부연 사이에 끼는 판재 | roof board | Panel inserted between the rafters or supports |
| 동해 | 외부 기온의 변화에 의하여 재료 내부에 함유된 수분의 동결 융해가 반복되면서 피해를 입게 되는 현상 | frost damage | The phenomenon where damage occurs due to the repeated freezing and thawing of moisture contained within the material, caused by changes in external temperature |
| 사주 | 하천 및 연안에 유사 등으로 인하여 생성된 퇴적지형 | sand bar | Sedimentary terrain formed by sedimentation due to tides and similar factors in rivers and coasts |
| 인상 | 중장비 등을 이용해 건설 자재를 들어올리는 일 | lifting | Lifting construction materials using heavy equipment |
| 취수 | 생활, 공업, 농업, 발전 및 하천유지용수 등을 위하여 하천이나 저수지의 물을 끌어오는 것 | water intake | Drawing water from rivers or reservoirs for living, industry, agriculture, power generation and river maintenance, etc. |

This domain-specific dataset includes multilingual translations of construction and shipbuilding industry terminology. Twenty-two languages are covered. The original term, which includes dialects, slang, or idiomatic usage, are taken from the Glossary of National Construction Standards, a publicly available dataset published by South Korea's Ministry of Land, Infrastructure and Transport (MOLIT), which defines standard construction terms used nationwide as well as the on-site interviews of workers. The translated term has been carefully reviewed with on-site workers and construction site managers to ensure relevant context is reflected.

The app features Speech-to-Text (STT) and Text-to-Speech (TTS) functions for accurate voice recognition in high-noise environments. Real-world tests using 1.05 million data samples confirmed over 90% translation accuracy and high user satisfaction. Users can also save frequently used technical terms for quick access, enhancing efficiency.

By minimizing translation errors, Safety Talk reduces safety risks and prevents industrial accidents. Additionally, it improves efficiency by ensuring clear work instructions, addressing the skilled labour shortage, and supporting aging workforces. As an AI-driven solution, Safety Talk is set to revolutionize communication, improve safety, and enhance productivity across construction sites worldwide.

**Partners**

- Crowdworks AI collaborated with a major South Korean construction firm (Company S) to co-develop an AI translation app tailored for the construction industry.
- Crowdworks AI' Role: To support the development of an AI translation app specialized in the construction domain, Crowdworks AI was responsible for building a construction-specific terminology dataset, designing metadata, optimizing the model through prompt engineering, fine-tuning, and conducting on-site tests in over 20 languages.

## 2.2 Benefits of the Use Case

- Industry, Innovation and Infrastructure
- The use case addresses a critical need for innovation in the construction sector. According to the Productivity Statistics Portal of the Korea Productivity Center (KPC) [2], South Korea's per capita labor productivity index declined in 2022... While productivity in most industries has improved, the construction sector has continued to fall behind, underscoring the need for technological interventions like AI-driven translation solutions. In this context, the development of an AI-powered interpretation app trained in construction-specific terminology represents a groundbreaking innovation in the sector.

By reducing language barriers, improving safety, and enhancing workflow clarity, this AI-powered app directly contributes to productivity gains and safer working environments—making it a valuable innovation for the construction and shipbuilding sectors.

| SDG 9 Requirements | Crowdworks AI's Solution |
|---|---|
| 1. Resilient Infrastructure | Enhances infrastructure resilience through improved communication and reduced accident risks |
| 2. Inclusive Industrialization | Facilitates seamless multilingual communication, promoting inclusivity for diverse workforces |
| 3. Sustainable Industrial Development | Provides a sustainable technological solution that continuously evolves with domain-specific datasets |
| 4. Innovation and Technological Advancement | Utilizes LLM, domain-specific dataset generation and in-context learning techniques (prompt engineering) for real-time translation solutions |
| 5. Improved Workplace Safety and Productivity | Significantly reduces misunderstandings, ensuring clearer safety protocols and higher operational efficiency |

## 2.3 Future Work

We will continuously collect newly emerging technical terms and on-site slang frequently used in construction sites to improve translation accuracy. In addition, we aim to further enhance the AI model to enable more accurate speech recognition even in noisy environments.

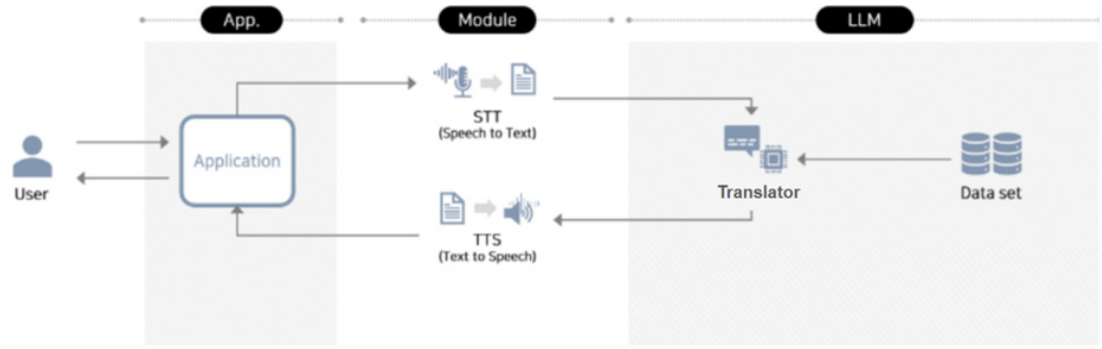
Furthermore, we plan to develop additional features that provide safety training and manuals for foreign workers, thereby contributing to the prevention of serious industrial accidents and promoting a culture of safety across the construction industry.

# 3 Use Case Requirements

- **REQ-01**: It is critical that effective communication for safety is ensured at construction sites with a high proportion of foreign workers. Language barriers in such environments can significantly impede communication and increase the likelihood of serious industrial accidents.
- **REQ-02**: It is critical to build and train AI models using domain-specific datasets that include construction-specific terminology and slang. Techniques such as fine-tuning and prompt engineering must be applied to improve translation accuracy.
- **REQ-03**: It is critical to validate the model's real-world applicability. Field testing was conducted in more than 20 languages, including English, Chinese, Hindi, Thai, and Vietnamese, achieving over 90% user satisfaction.

- **REQ-04**: It is critical that the app performs reliably in noisy construction environments. To this end, STT (Speech-to-Text) and TTS (Text-to-Speech) capabilities were incorporated to enhance voice recognition accuracy and improve user experience.

## 4 Sequence Diagram




## 5 References

[1] Ministry of Land, Infrastructure and Transport (MOLIT), "Glossary of National Construction Standards." [Online]. Available: https://www.molit.go.kr/USR/BORD0201/m_72/DTL.jsp?id=eng_0101&cate=&mode=view&idx=308987

[2] Korea Productivity Center, "Productivity Statistics Portal," 2023. [Online]. Available: https://www.kpc.or.kr/eng/statistics

[3] Google Research, "Gemma 2 Models - GitHub Repository," 2024. [Online]. Available: https://github.com/google/gemma

# Use Case 10: FAHR Smart App - AI-powered HR Solutions for Government Employees

الهيئة الاتحادية للموارد البشرية الحكومية
Federal Authority For Government Human Resources

**Organization**: Federal Authority for Government Human Resources (FAHR)

**Country**: UAE

**Contact Person**:

Samer Al Rahal, awards@fahr.gov.ae

## 1 Use Case Summary Table

| Item | Details |
|---|---|
| Category | Digital Government, AI for Human Capital Management |
| Problem Addressed | Manual HR processes across UAE federal entities are time-consuming, inconsistent, and not personalized. There's a need for a unified, AI-enhanced HR system to support workforce efficiency and decision-making. |
| Key Aspects of Solution | – AI-powered digital assistant for instant HR support<br>– Automated leave and attendance workflows<br>– Smart learning and career path recommendations<br>– Real-time analytics dashboards<br>– Seamless integration with federal systems |
| Technology Keywords | Generative AI (GenAI), AI Agents, Natural Language Processing (NLP), React Native, Next.js, AI Chatbot, Predictive Analytics |
| Data Availability | Private |
| Metadata (Type of Data) | Text, User Interaction Logs, Behavioral Analytics, Structured HR Data |
| Model Training and Fine-Tuning | Generative AI fine-tuned using secure, anonymized HR datasets; NLP and classification models customized for UAE federal HR terminology and processes |
| Testbeds or Pilot Deployments | [FAHR Smart App deployed across all UAE federal entities - Phase I launched April 2024] More details are available on the official FAHR website. [Online]. Available: https://www.fahr.gov.ae/en/smart-app-launch.aspx[4].<br>[User Experience (UX) labs were conducted during 2023–2024 with integration of user feedback from federal stakeholders. A summary report can be found here:[Online]. Available: https://www.fahr.gov.ae/en/ux-feedback-report.aspx[5]. |
| Code repositories | Private - managed internally under UAE Federal infrastructure (can be shared on request if applicable) |

# 2 Use Case Description

## 2.1 Description

The Federal Authority for Government Human Resources (FAHR) Smart App leverages Generative Artificial Intelligence (AI) to enhance Human Resources (HR) services for federal government employees [1]. It provides AI-driven insights, automated leave and attendance management, smart learning recommendations, and AI chatbots for real-time support [2]. The app aligns with the UAE Digital Government Strategy [3], contributing to a more sustainable and efficient public workforce.

### Partners

| Partner | Role |
|---|---|
| UAE Federal Government Entities | End-users and beneficiaries of smart HR services |
| Telecommunications and Digital Government Regulatory Authority (TDRA) | Oversight and digital integration |
| FAHR AI and Innovation Teams | Development and operationalization of AI features |

## 2.2 Benefits of the Use Case

Justify selection by linking specific project outcomes with the targets of these global goals:

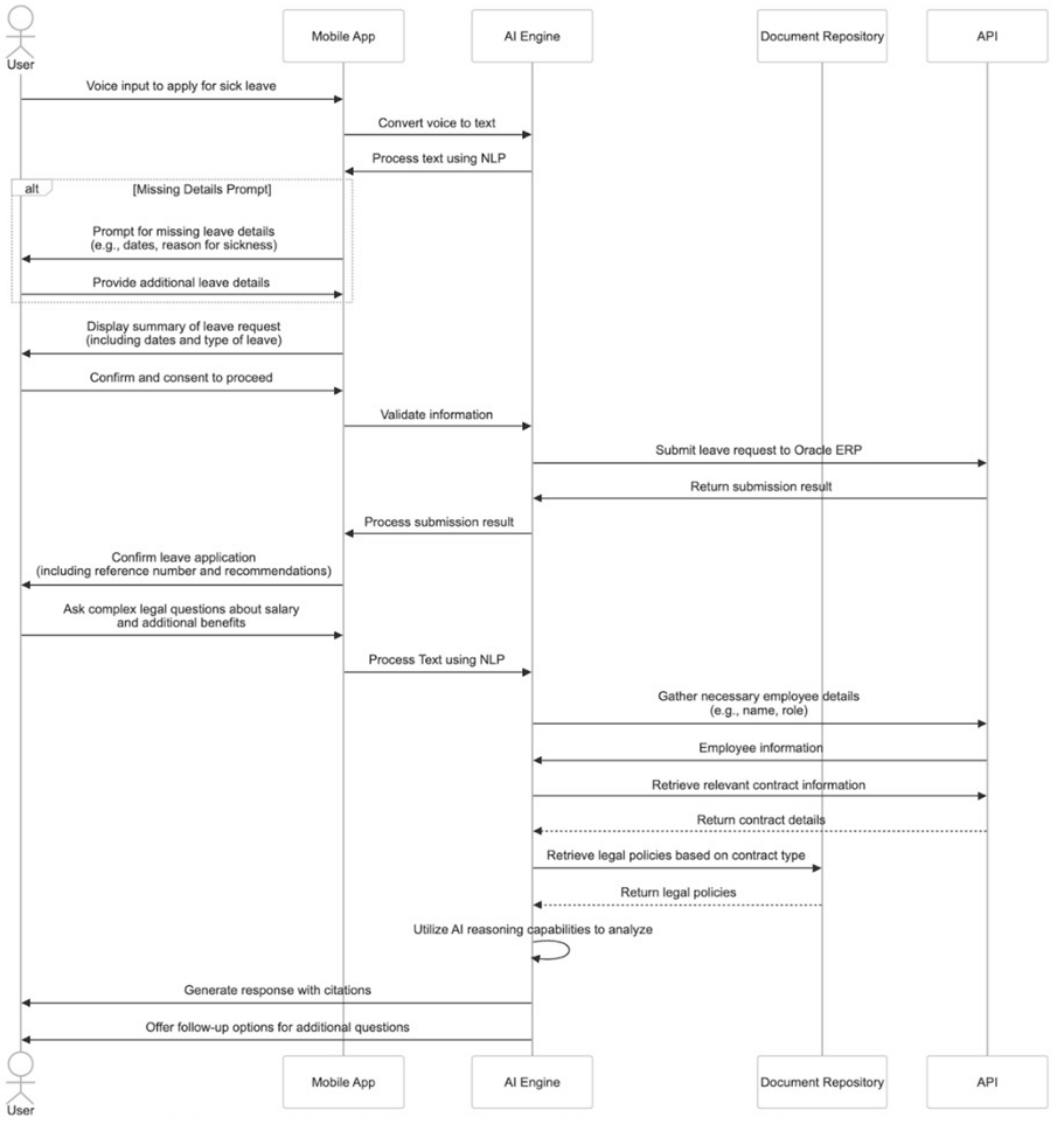
- The FAHR Smart App supports sustainable economic growth by optimizing HR processes, improving workforce efficiency, and ensuring better job satisfaction.
- By incorporating AI-driven HR solutions, the app fosters technological innovation and enhances digital infrastructure. The FAHR Smart App uses cutting-edge technology such as Generative AI (GenAI), AI Agentic framework, React, Next.js, etc.
- The app promotes strong institutional frameworks through transparent and data-driven decision-making for HR policies in government entities. It reduces reliance on manual interventions using AI-automated solutions like the AI digital assistant.

## 2.3 Future Work

Future work will focus on enhancing AI capabilities in the FAHR Smart App by integrating predictive analytics for workforce planning, personalized career development recommendations, and AI-driven employee engagement tools. Further, establishing AI governance frameworks and compliance with International Telecommunication Union (ITU) standards will be a priority. Collaboration with international AI research institutions will help refine models and ensure ethical AI usage.

## 3 Use Case Requirements

## 4 Sequence Diagram




## 5 References

[1] [1] Tom B. Brown et al., "Language Models are Few-Shot Learners" (OpenAI, 2020), accessed June 24, 2025, https://arxiv.org/abs/2005.14165.

[2] Accenture, "AI in Human Resources: Enhancing Workforce Management," White Paper (2023).

[3] UAE Government, "Digital Government Strategy 2025," TDRA (2022), accessed June 24, 2025, https://tdra.gov.ae/digital-strategy.

[4] "Smart App Launch," Federal Authority for Government Human Resources (FAHR), accessed June 24, 2025, https://www.fahr.gov.ae/en/smart-app-launch.aspx.

[5] "UX Feedback Report," Federal Authority for Government Human Resources (FAHR), accessed June 24, 2025, https://www.fahr.gov.ae/en/ux-feedback-report.aspx.

# Use Case 11: Virtual Legal Assistant of the Rights Mapping and Analysis Platform (RiMAP) to Improve Legal Analysis and Access to Rights and Solutions for Forcibly Displaced and Stateless Persons

**Organization:** United Nations High Commissioner for Refugees (UNHCR)

**Country:** Switzerland (scope is multi countries)

**Contact Person(s)**:

Esther Benizri: benizri@unhcr.org;

Moises Maldonado: maldonam@unhcr.org.

## 1 Use Case Summary Table

| Item | Details |
|---|---|
| Category | Intelligence system / Humanitarian |
| Problem Addressed | Difficulty collecting, analyzing, and maintaining up-to-date legal information affecting forcibly displaced and stateless persons across numerous countries, hindering effective legal support and advocacy. |
| Key aspects of Technology | 1. Automated legal data collection & updating from diverse sources.<br>2. AI-powered analysis (summarization, extraction). Solution<br>3. Human-in-the-loop review for quality assurance.<br>4. Internal VLA interface & future external chatbot. |
| Data Keywords | Large Language Models (LLMs), Natural Language Processing (NLP), Technology, AI Chatbot, Web Scraping/Crawling, Data Extraction, Summarization, Text Generation |
| Data Availability | First private, later public (https://rimap.unhcr.org/), composed of domestic laws and policies regulating access to rights for forcibly displaced and stateless persons. |
| Metadata | Legal texts (laws, regulations, policies), case law documents, (Type narrative legal and factual reports, government publications, and of Data)<br>UNHCR guidance documents. |
| Model Training and Fine Tuning | Utilizing localized pre-trained LLMs fine-tuned on relevant legal corporation and UNHCR-specific documents for summarization, question and answering, chat with documents, and data extraction. |
| Deployments | Internal Prototype evaluated by UNHCR staff. |
| Code | To be shared publicly on the UN AI Toolbox repositories. |

# 2 Use Case Description

## 2.1 Description

### Background

The number of individuals forcibly displaced from their homes—whether to another country or within their own borders—continues to rise, driven by factors such as war, violence, persecution, and climate change. Given the recent global events, it is unfortunately unlikely that this trend will reverse in the near future. These populations often face prolonged displacement, spending years or even decades away from their homes, with prospects for return growing increasingly uncertain as time passes.

At the same time, the forcibly displaced are expected to gradually integrate into their host countries or communities. However, legal frameworks and programs are not always adequately equipped to support their potential. Gaps in national laws can hinder access to rights and create challenges in implementing policies designed to assist them.

UNHCR's core mandate is to protect the rights of forcibly displaced and stateless persons. Through legal analysis, UNHCR and its partners can identify gaps or inconsistencies in legal frameworks and advocate for legal and policy reforms that enhance protection and support the rights of displaced and stateless populations.

Legal analysis is critical for understanding how national legal systems can better integrate forcibly displaced and stateless persons into host countries. It allows for the identification of legal barriers to protection and social and economic inclusion, such as restrictions on employment, education, or access to healthcare. By analyzing the legal rights and protections afforded to displaced and stateless persons, UNHCR can work with governments to develop policies that promote protection and sustainable solutions, reducing the risks of marginalization and social exclusion.

Legal analysis provides the evidence needed to support advocacy for legislative reforms at national, regional and international levels. Gaps in the protection frameworks could include lack of access to asylum, or limited access to documentation. By providing a comprehensive analysis of these issues, UNHCR can push for legal reforms that align with international human rights standards.

UNHCR developed the Rights Mapping and Analysis Platform (RiMAP), a publicly available web-based initiative designed to centralize comprehensive legal and policy information related to forcibly displaced and stateless populations. While RiMAP serves as a critical resource, collecting and maintaining up-to-date information across multiple countries remains a significant challenge due to limited human and financial resources.

Often, UNHCR faces challenges in collecting legal information, as national legal frameworks and resources are often fragmented across various legal libraries and platforms but also because staff simply lack the time, faced with multiple priorities, to comprehensively research and analysis country level legal frameworks affecting core rights. Given the constraints on staff time and resources, UNHCR is partnering with academia and NGOs to support the collection of legal data. However, the scale of this effort is tremendous. Ensuring the long-term sustainability

and reliability of RiMAP is dependent on the ability to continuously update its content in a timely and efficient manner.

## Objectives and aims of the use case

This use case forms part of the broader development of the Rights Mapping and Analysis Platform (RiMAP). As part of the ongoing efforts to enhance RiMAP, UNHCR is integrating the internal Virtual Legal Assistant (VLA) powered by Large Language Models (LLMs) to facilitate legal data collection, improve policy analysis, and make this information more accessible to internal users, focusing on data translation, extraction summarization as well as production of structured outputs.

Primary Objective (Internal Focus): To support UNHCR country editors and legal advisors by developing and integrating an internal Virtual Legal Assistant (VLA) powered by LLMs. The VLA will process and analyze large volumes of legal documents from national sources across all UN countries/territories, facilitating legal data collection (extraction, summarization) and improving policy analysis efficiency. Target: Aim to reduce preliminary legal research time for targeted tasks by 30% within 12 months of full deployment.

Once data is collected, the AI assistant is able to extract key legal insights and summarize complex legal texts. This includes government publications, legal databases, academic research, and other publicly available resources related to the legal rights and status of displaced and stateless persons. By leveraging AI, the VLA will enable UNHCR to analyze evolving legal landscapes, ensuring that the organization stays up to date with legal developments. By automating this process, the VLA will greatly reduce the manual effort required to track and gather legal data, ensuring that the platform remains current without the heavy workload traditionally involved in such research.

The secondary goal of this use case is to expand the VLA's functionality to support users. By transforming the VLA into an AI-powered chatbot, the system provides easily accessible and user-friendly legal information derived from the RiMAP library. This will be a significant step towards making legal data more accessible within UNHCR but importantly also to a wide range of stakeholders, including universities, legal aid organizations, governments, development actors, civil society groups, and, importantly, forcibly displaced and stateless persons themselves.

The anthropomorphic chatbot interface offers an intuitive interaction that allows users to interact with the VLA with natural language, ask questions, and receive legal information in a more conversational and accessible format. The VLA's capabilities will span across all UN member states and territories, ensuring that users have access to a comprehensive, global legal resource. This will empower users to navigate complex legal systems and better understand the rights of displaced and stateless persons, enabling more informed advocacy and facilitating greater legal support for such populations.

## Innovative technological approach implemented

The VLA automatically scrapes, extracts, and categorizes relevant legal documents from these diverse sources, allowing it to build and expand its repository of up-to-date legal information continually. The VLA automates the extraction and summarization of relevant legal data from vast volumes of legal text corpora, policy documents, and public legal databases.

As new legal documents are acquired or existing ones updated, the system automatically index or re-index them, ensuring that the underlying data used for analysis and search remains current and readily accessible

To ensure the relevance and accuracy of the information in RiMAP, the system relies on a robust process of human review and selection. While AI automates the initial data collection, human experts review and select the most relevant and up-to-date information that serve to produce content. This hybrid approach combines the speed and efficiency of AI with the accuracy and contextual understanding provided by human oversight, ensuring that only the most pertinent legal data is included in the platform.

Once users pre-select legal documentation and interact with the Assistant, they evaluate the output against four criteria (acceptability, correctness, relevance and appropriateness). The evaluation of the user's input is used to re-instruct (prompt) the Assistant to improve the expected output, in a continuous optimization loop, until the user is satisfied with the response.

By linking to external public legal databases, the VLA ensures that the information within RiMAP remains continuously updated, allowing users to access the most current and reliable legal resources.

### Impact and benefits of the solution

The VLA will significantly reduce the time and effort required for legal research and analysis by automating the collection and summarization of information from multiple external sources. By centralizing and indexing legal data, the VLA will allow users to access accurate and up-to-date legal knowledge quickly and maintain the quality of the legal analysis, as all the text generated is referenced to the original documents' official source.

The human review process integrated into the VLA will enhance the relevance and quality of the information, ensuring that RiMAP's content reflects the most reliable and accurate legal data available. This approach will also reduce the time UNHCR staff spend on maintaining the platform, improving operational efficiency.

The VLA chatbot will offer users an intuitive interface that allows them to interact with the VLA, ask questions, and receive legal information in a more conversational and accessible format. The VLA's capabilities will span all 193 UN member states, ensuring users have access to a comprehensive, global legal resource. This will empower users to navigate complex legal systems and better understand the rights of displaced and stateless persons, enabling more informed advocacy and facilitating greater legal support for displaced populations.

### Partners

UNHCR published a Call for Expression of Interest to gradually build a legal network of partners that will contribute to the review and analysis of the legal data collected for RiMAP. They will be using the VLA to enhance efficiency to produce legal analysis.

## 2.2 Benefits of the Use Case

The combination of AI-driven technology and human oversight will enable RiMAP to remain a comprehensive, accurate, and continuously updated resource, supporting global efforts to

reduce inequality, promote justice, and ensure the protection of displaced persons' rights. The use case will contribute to achieving SDG Goals 10 (Reduced Inequality) and 16 (Peace, Justice, and Strong Institutions), ultimately empowering displaced populations by improving their access to legal rights and protections.

## 2.3 Future Work

The immediate next step is the development of an internal VLA prototype to support editors, following which the VLA will be expanded to serve as a full-scale content generating tool, focusing on the core functionalities, such as summarizing and researching large legal documents, extracting legal data, and providing accurate legal analysis. The VLA will be tested internally by UNHCR staff to validate the AI-driven features and to ensure the system can meet the practical needs of legal research and policy analysis. Feedback from this phase will inform the refinement and scaling of the VLA.

One of the key future goals is to fully automate the legal search and regular updates. Currently, legal research often requires significant substantive manual input and time. In future iterations, the VLA will leverage AI to automatically update RiMAP with relevant information pulled from legal sites, making the process faster and more efficient. Automation will extend to continuous data collection and updates, ensuring that the information in RiMAP is always current and relevant.

Future enhancement will be the design of a public AI-powered chatbot for external users, including universities, legal aid organizations, governments, civil society, and forcibly displaced and stateless persons. This chatbot will serve as an intuitive interface, allowing users to ask legal questions, search for relevant information, and receive clear, summarized legal guidance directly from RiMAP's library. The chatbot will be designed for ease of use, enabling external stakeholders with varying levels of legal knowledge to access relevant and up-to-date legal information quickly and efficiently.

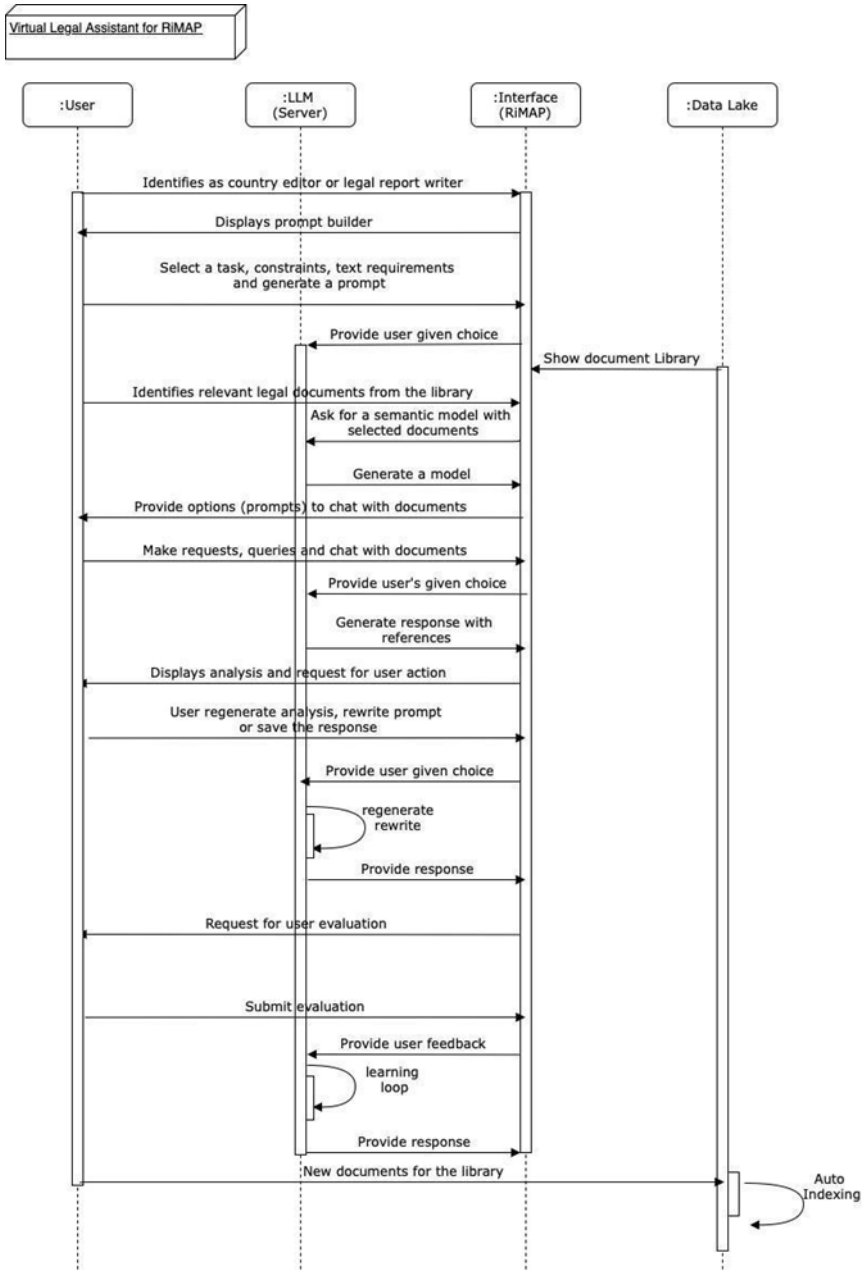
Finally, future work will focus on extending language support to ensure broader accessibility to legal information across regions. This will include more languages with specific legal terminologies relevant to various geographical areas, enabling the VLA to cater to a wider and more diverse user base.

# 3 Use Case Requirements

- **REQ-01**: Functional: The VLA shall summarize legal documents (up to 100 pages) into 5-15 key paragraphs with verifiable citations to the source document.
- **REQ-02**: Functional: The VLA shall extract predefined data points from designated legal text types.
- **REQ-03**: Functional: The VLA chatbot shall respond to user queries using *only* verified information in the RiMAP database, providing source links.
- **REQ-04**: Functional: The system shall automatically trigger a document's reindexing upon detection of an update in its source or upon manual upload of a newer version.
- **REQ-05**: Non-Functional: Summarization requests shall be processed with an average response time under 60 seconds.
- **REQ-06**: Non-Functional: The system must maintain an audit trail of AI suggestions and human review actions.

- **REQ-07**: Operational: The human review interface must allow editors to easily accept, reject, edit, and comment on AI-generated outputs.
- **REQ-08**: Data: The VLA must be configured to prioritize and process legal documents from specific, authenticated sources defined by administrators.

## 4 Sequence Diagram




## 5 References

[1] Cheatham, Susan. "UNHCR High Level Business Requirements." Internal UNHCR report.

[2] UNHCR. "RIMAP." Accessed June 24, 2025. https://rimap.unhcr.org/.

# Use Case 12 : Implementation Framework for AI-Enabled HRM in UAE Limited Liability Companies

**Organization**: Dubai Government Human Resources Department; Policies and Programs Sector - Human Resources Information Department

**Country**: United Arab Emirates

**Contact person(s)**:

Abdulla Abdulaziz: Alanjawi@dghr.gov.ae,

Soha Jawabreh: sjawabreh@dghr.gov.ae

## 1 Use Case Summary Table

| Item | Details |
|---|---|
| Category | Productivity |
| Problem Addressed | There is a critical lack of structured, evidence-based tools to help organizations implement and evaluate AI in human resource management. Without clear guidance, companies struggle to understand how AI can reduce hiring costs, improve decision-making, and deliver measurable value especially in cost-sensitive business environments. |
| Key Aspects of Solution | • Decision-making toolkit for AI in Human Resource Management(HRM).<br>• Evaluation framework for AI effectiveness in cost reduction.<br>• Strategic recommendations for AI integration in the UAE business context. |
| Technology Keywords | Artificial Intelligence, Human Capital Management, HR Analytics, Return on Investment(ROI) Assessment, AI Implementation Framework |
| Data Availability | Private – Includes survey data, interview data, cost metrics, and ROI assessments from participating UAE Limited Liability Company(LLCs). |
| Metadata (Type of Data) | Textual survey responses, interview transcripts, numerical cost metrics, and ROI figures. |
| Model Training and Fine-Tuning | Approaches involve quantitative analysis of survey data and qualitative coding of interview data for toolkit development. |
| Testbeds or Pilot Deployments | Pilot deployments in process with participating UAE LLCs for empirical validation. |

# 2 Use Case Description

## 2.1 Description

The increasing interest in Artificial Intelligence (AI) within HRM has created a pressing need for practical tools that facilitate effective implementation particularly in the context of LLCs in the UAE. Despite the rapid digital transformation, many HR departments struggle with identifying cost-effective ways to integrate AI technologies due to limited strategic frameworks and actionable guidance.

This Doctor of Business Administration(DBA) research addresses this gap by developing a comprehensive decision-making toolkit tailored to HR professionals in UAE-based LLCs. The study introduces a practical framework to assess and measure AI's impact on reducing HRM-related costs, while also offering strategic recommendations for AI integration. Technically, innovation lies in translating complex AI principles into user-friendly tools that support operational decisions. Practically, the toolkit enables HR teams to streamline processes, enhance data-driven decision-making, and boost organizational efficiency contributing to long-term sustainability and agility in a fast-evolving business landscape.

Objectives and Aims:

- To develop a decision-making toolkit for HR managers to guide AI implementation with a focus on minimizing costs and maximizing ROI.
- To establish a framework for evaluating the effectiveness of AI tools in reducing HRM-related costs in UAE LLCs.
- To provide strategic recommendations tailored to the UAE business environment for effective AI integration in HRM operations.
- To empower HR professionals in making data-driven decisions and improving operational efficiency through AI.

Innovative technological approach being implemented.

- The innovative approach of this use case lies in the development and deployment of a comprehensive AI-driven toolkit tailored for HRM in UAE LLCs. It combines quantitative analytics (from surveys of HR professionals) and qualitative insights (from interviews with senior HR managers) to generate data-informed decision-making tools.
- The solution introduces a structured, locally relevant AI implementation framework - not only guiding organizations through the adoption process but also enabling them to measure the ROI and cost-effectiveness of AI tools in HRM.
- Furthermore, the project integrates plans for future automation of these frameworks into a web-based or mobile platform, making AI adoption more accessible, scalable, and impactful across the sector.

**Use Case Status:** Advanced Doctoral research stage.

**Partners:** Dubai Government Human Resource Department, University of Derby UK

## 2.2 Benefits of the Use Case

The use case will lead to better economic growth, reduced inequality in addition to enhancing workplace efficiency and innovation through AI integration in HRM practices.

It also contributes to strengthening technological capabilities in HR operations using AI for mitigating biases in hiring and talent management.

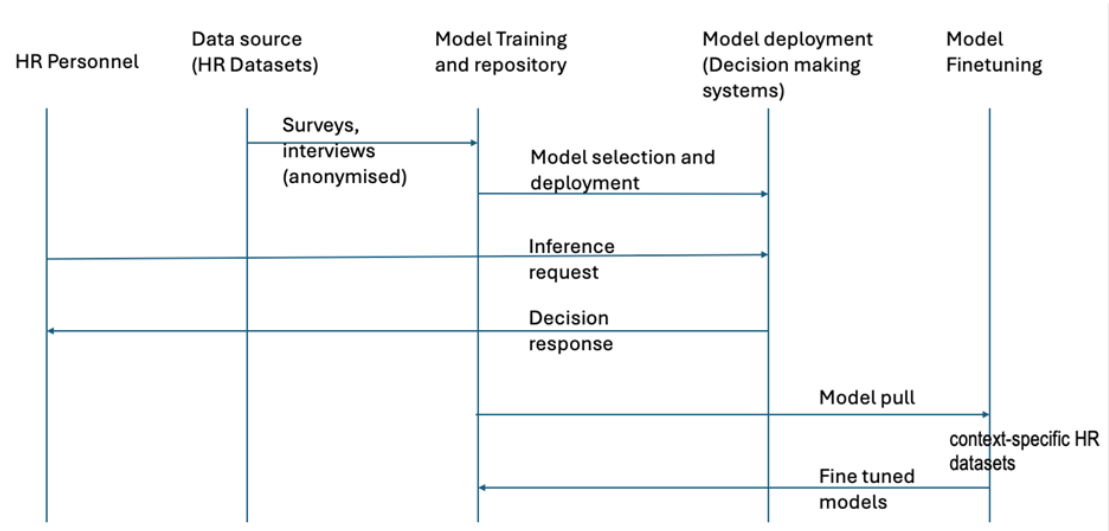
## 2.3 Future Work

- Collaborations with tech providers and AI solution vendors to integrate cutting-edge tools into the HRM framework.
- Expansion to other sectors (e.g., healthcare, education, logistics) to apply the AI decision-making toolkit beyond HR.
- Partnerships with government innovation entities and academic institutions to support further research, deployment, and policy alignment.

# 3 Use Case Requirements

- it is critical that a comprehensive decision-making toolkit for HR managers with integration of AI in HRM is available for optimizing decision in HR.
- it is critical that a structured framework/tool for measuring AI tools' effectiveness in reducing HRM costs is implemented.
- it is critical that strategic recommendations for optimizing AI integration specific to the stakeholder's business context. This should be based on a documented gaps in current practice where organizations lack practical tools for implementing and evaluating AI solutions in HRM operations.

# 4 Sequence Diagram




# 5 References

[1] University of Derby, "Postgraduate Research Project Login," accessed June 24, 2025, https://research.derby.ac.uk/do/derby-authentication/login?rdr=/9qy4z/postgraduate-research-project.

# Use Case 13: Lost in Translation? AI's Promise and Peril for Media in Low-Resource Languages

**Organization**: Fondation Hirondelle, Media for peace and human dignity

**Country**: Switzerland

**Contact person**:

Caroline Vuillemin, caroline.vuillemin@hirondelle.org, +41 21 654 2017

## 1 Use Case Summary Table

| Item | Details |
|---|---|
| Category | Productivity |
| Problem Addressed | The case tackles the critical issue of equitable access to Artificial Intelligence (AI) tools for media organizations worldwide, in Eastern DRC in particular. As AI rapidly transforms journalism and the information ecosystem, a significant digital divide is emerging, leaving media outlets using low-resource languages at a disadvantage. |
| Key Aspects of Solution | The case brings together media practitioners and technology to explore the challenges faced by these media, with the examples of DRC, highlighting the lack of data in Swahili, suitable algorithms, and technical resources. We will showcase innovative solutions, particularly from African-led initiatives, and discuss the potential for AI to empower these outlets, focusing on automated transcription and translation |
| Technology Keyword | AI based translation and transcription |
| Data Availability | public |
| Meta Data | audio samples |
| Model Training and Fine-tuning | N/A |
| Testbeds of Pilot Deployment | N/A |
| Code Repositories | N/A |

# 2 Use Case Description

## 2.1 Description

Use-case description:The case tackles the critical issue of equitable access to Artificial Intelligence (AI) tools for media organizations worldwide, in Eastern DRC in particular. As AI rapidly transforms journalism and the information ecosystem, a significant digital divide is emerging, leaving media outlets using low-resource languages at a disadvantage. The case brings together media practitioners and technology to explore the challenges faced by these media, with the examples of DRC, highlighting the lack of data in Swahili, suitable algorithms, and technical resources. We will showcase innovative solutions, particularly from African-led initiatives, and discuss the potential for AI to empower these outlets, focusing on automated transcription and translation. What does African media need to be able to leverage this transformative technology? How can we ensure that the rich diversity of languages, cultures, and local knowledge are embodied in the process? Our goal is to foster dialogue and collaboration, urging greater investment and support for a more inclusive media landscape.

**Partner**

Prof. Isabel Günther - Development Economics - ETH Zurich - Clausiussstrasse 47, 8092 Zürich
Partner url: https://eth4d.ethz.ch/

## 2.2 Benefits of the Use Case

Access to information and the role of media play a key role to secure peace and justice in any society. Yet, when local journalists from low resource languages and with limited access to new technology can't work in the language the majority of their audience speaks, then there is an inequality, and a gap. Bridging that gap for more peaceful and just societies is important and can be done if local media in developing countries can access new tools and if local populations can access information in the language they speak and in the media they locally use.

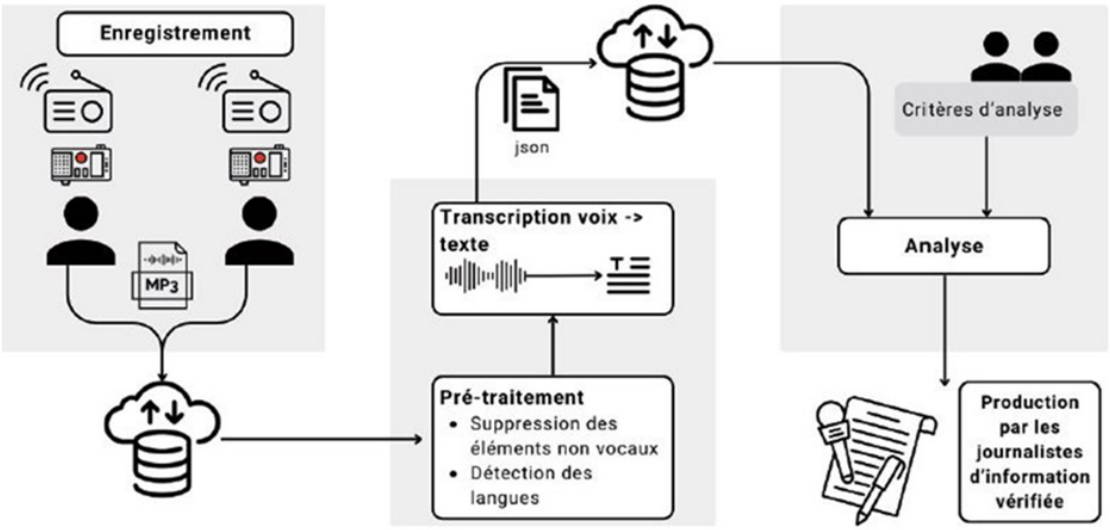
## 2.3 Future Work

Data collection, Create new variations/extensions to the same use case, Standards development related to the use case Elaborate proposal: The case is currently monitoring 10 radio stations in Eastern DRC, in Swahili and French. The transcripts of the content is done by AI, and the search for disinformation or problematic content such as lies or hate speech is done by the machine based on key words and knowledge provided by the local journalists.

We need to improve the speech to text in Swahili with more radio stations' contents, as well as constantly update the trends to monitor as the situation in Eastern DRC is volatile and fragile.

# 3 Use Case Requirements

## 4 Sequence Diagram




## 5 References

# Use Case 14: AI-Driven Wireless Audio Conference System for Smartphone

**Organization**: Mbeya University of Science and Technology (MUST)

**Country**: Tanzania

**Contact Person**:

Jackson Braiton Arone, jacksonnkizingi@gmail.com

## 1 Use Case Summary Table

| Items | Details |
|---|---|
| **Category** | Productivity |
| **Problem to be Addressed** | In today's event venues, typical wireless microphone systems predominantly use Frequency Modulation (FM)radio frequencies. Unfortunately, these systems often face various problems such as a lack of available microphones, interference, weak signal quality, overlapping speech, slow turn-taking, and manual processing of speaking request. So, prioritize speakers, handle requests, and guarantee uninterrupted, equitable, and seamless communication, a clever system is required. An automated, well-organized, and equitable communication flow is required. |
| **Key Aspects of Solution** | The proposed solution leverages a smartphone app and a Personal Computer(PC)-based server orchestrated entirely by an AI—eliminating manual intervention—to manage speaker requests, prioritize talk turns, and process audio streams in real time.<br><br>The AI uses pre-established rules (such as first-come-first-served, user roles, and speaking history) to give or queue access when a user hits "Talk Request." The server then instantly transmits the request to the AI. After activating, the AI continuously checks the audio for format compatibility and noise reduction before streaming it to the PA system. It then immediately releases the slot and advances the next user, AI-driven must apply noise reduction, echo cancellation, and voice activity detection to clean up audio before it's sent to the Public Address(PA)system, ensuring clear and professional audio quality even in noisy environment by the aid of Neural networks + Digital Signal Processing(DSP), where DSP is a traditionally method. |
| **Technology Keywords** | AI, Wi-Fi Communication, Mobile App, Computer App and Audio System |
| **Data Availability** | Public (Limited real-time user and audio data during system use) |

(continued)

| Items | Details |
|---|---|
| **Metadata (Type of Data)** | Videos data |
| **Model Training and Fine-Tuning** | Past speaker behaviour patterns to predict mic abuse or speech quality.<br><br>Neural networks + DSP ( Audio classification e.g., noise vs. speech).<br><br>Natural Language Processing(NLP)for understanding content of speech (for moderation).<br><br>Queue optimization using reinforcement learning or priority policies. |
| **Testbeds or Pilot Deployments** | In academic and public service environments to validate its effectiveness and usability. Also local town hall meetings, where the system was integrated with public address (PA) systems to facilitate transparent and structured civic engagement. |

# 2 Use Case Description

## 2.1 Description

Introduction: In many professional and academic settings, such as conferences, meetings, and large gatherings, managing real-time communication and ensuring clear audio for all participants is a significant challenge. Traditional microphone systems are often limited, immobile, or insufficient to meet the needs of large groups, resulting in inefficiencies, delays, and potential disruptions. Participants may struggle to access microphones, leading to frustration and reduced engagement. Additionally, the responsibility for managing these systems often falls to IT departments, who may lack the necessary expertise in audio management. With the rise of smartphones and wireless communication, there is an opportunity to utilize existing technology to provide a more seamless and efficient communication system.

Solution Overview: The solution integrates AI with a smartphone app and a PC-based server, creating a fully automated speaker management system. When a user requests to speak, the server communicates with the AI module, which decides whether to grant access or add the user to a queue based on factors like request order, user roles, and speaking history. The system also continuously processes the audio stream, applying noise reduction and ensuring compatibility with the PA system. This innovative approach leverages real-time multimedia applications (RTMA), Wi-Fi connectivity, and AI decision-making, automating what was once a manual process and making communication more efficient and accessible.

The expected impact includes increased efficiency by automating speaker management, enhanced user experience with smartphones as microphones, and reduced costs by eliminating the need for fixed audio equipment. This ensures smoother communication, improves productivity, and maintains audience engagement.

### Partner

Mbeya University of Science and Technology (MUST)

## 2.2 Benefits of the Use Case

This system improves access to spoken communication in educational settings, offering real-time transcription and equitable speaker turns. It ensures all students can participate meaningfully, even when traditional audio tools are unavailable or limited.

By enforcing a fair, AI-managed speaking queue, the solution prevents conversational dominance and ensures all voices—including those from underrepresented genders—have equal opportunity to be heard.

The system reduces costs related to specialized audio hardware and technical staffing. It provides a scalable communication tool for resource-limited workplaces, supporting productivity and innovation, especially in developing regions.

Replacing traditional hardware with AI-powered, software-based infrastructure supports digital transformation. Using existing smartphones and local Wi-Fi makes the solution deployable even in areas with limited infrastructure.

The AI ensures fair speaker selection regardless of background. Features like transcription aid those with hearing impairments or softer voices, helping everyone participate equally.

This system improves communication in public forums, town halls, and councils by promoting organized participation and public engagement. It also aids in transparent decision-making.

Automated recording, transcription, and fair speaking procedures promote transparency and accountability in official settings. This supports stronger governance and civic trust.

## 2.3 Future Work

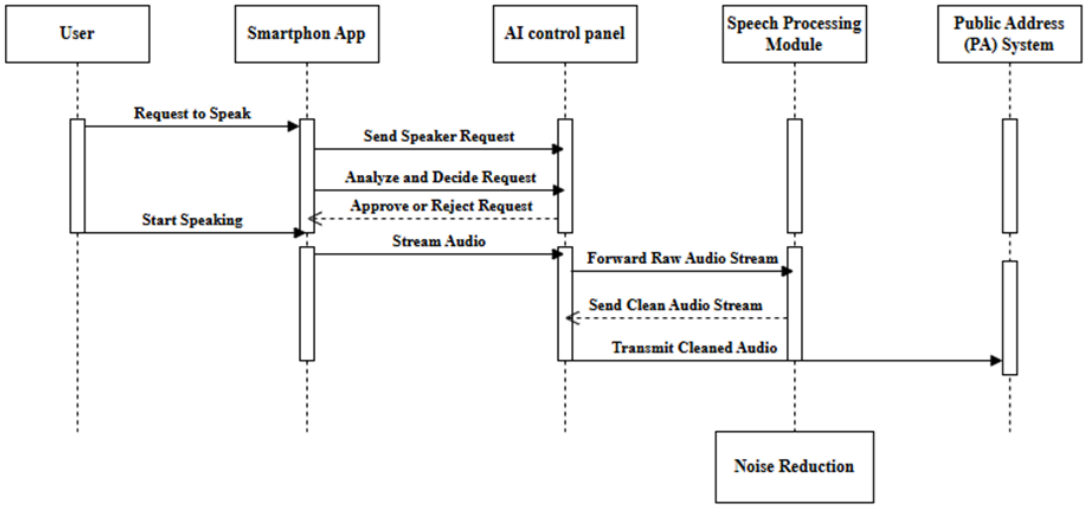
As the AI-Driven Wireless Audio Conference System evolves, future work will focus on enhancing AI capabilities through features like dynamic speaker prioritization, sentiment-based urgency detection, and speaker diarization for better multi-speaker management. The system will expand to support cloud-based deployments for remote and hybrid meetings, and multilingual transcription and translation to promote inclusivity. Technical upgrades will include support for low-spec devices, integration with Bluetooth PA systems, and edge computing to reduce latency. Real-world audio data will be collected to retrain and fine-tune models for improved accuracy across diverse environments and accents. Strategic collaborations with universities, local governments, and accessibility experts are planned to expand reach and ensure inclusive design. Lastly, development will emphasize strong data privacy, ethical AI governance, and compliance with international standards like General Data Protection Regulation (GDPR).

# 3 Use Case Requirements

- **REQ-01**: It is critical that the mobile application must be compatible with Android and iOS platforms to ensures that users, regardless of the device they use, can join and participate in meetings, conference, maximizing inclusivity and adoption.
- **REQ-02**: It is critical for all mobile devices (clients) and the AI-driven server to be connect to the same local Wi-Fi network. This setup avoids internet latency, ensures fast, low-latency communication, and simplifies deployment in venues like schools, conferences, or town halls.

- **REQ-03**: It is crucial that the AI-driven server must analyse incoming speaker requests, evaluate parameters like user role, speaking time, and order of request, and automatically approve, reject, or queue them based on predefined rules. This enables fair, organized, and efficient turn-taking.
- **REQ-04**: To ensure clear and professional audio quality, even in noisy environments, the AI-driven system must apply noise reduction, echo cancellation, and voice activity detection before sending audio to the PA system.
- **REQ-05**: The system must handle many users at once. The server must be able to track active, waiting, and muted users, and switch audio sources seamlessly between speakers to avoid overlaps, dead air, or audio collisions.

## 4 Sequence Diagram




## 5 Reference

[1] A. V. Oppenheim and R. W. Schafer, "Discrete-Time Signal Processing," 3rd ed., Pearson, 2009.

[2] D. Jurafsky and J. H. Martin, "Speech and Language Processing," 3rd ed. (draft), 2023. [Online]. Available: https://web.stanford.edu/~jurafsky/slp3/

[3] R. S. Sutton and A. G. Barto, "Reinforcement Learning: An Introduction," 2nd ed., MIT Press, 2018.

# Use Case 15: AI-Powered Universal Communication Platform: Breaking Down Barriers Through Collaborative Innovation

**Country**: United Kingdom

**Organization**: The Scott-Morgan Foundation

**Contact Person**:

LaVonne Roberts, lavonne@scottmorganfoundation.org, +1 512 695-4500

## 1 Use Case Summary Table

| Item | Details |
|---|---|
| Category | Productivity |
| Problem Addressed | Over 500 million people worldwide face barriers to communication. Research shows only 18.2% of students needing communication support receive any form of Augmentative and Alternative Communication(AAC), with just 4.8% accessing speech-generating devices. Traditional assistive technologies exist in isolated silos-prohibitively expensive, medically classified, and lacking integration. |
| Key Aspects of Solution | Eye-tracking and adaptive input systems with Large Language Models (LLMs) for prediction Personalized voice synthesis (ElevenLabs) and avatar representation (D-ID) Open ecosystem approach connecting previously siloed technologies Consumer hardware platform reducing costs by 80% |
| Technology Keywords | Ecosystem integration,Voice synthesis,Avatar generation, Eye- tracking, LLMs Assistive technology |
| Data Availability | Private - Personal voice samples and user interaction data (with strict privacy controls) |
| Metadata (Type of Data) | Audio (voice samples), Image (eye-tracking data,facial expressions), Text (user communication) |
| Model Training | Voice models from limited samples LLMs for prediction Computer vision for eye-tracking |

(continued)

| Item | Details |
|---|---|
| Testbeds or Pilot Deployments | Successfully demonstrated with Amyotrophic lateral sclerosis (ALS) users at the Consumer Electronics Show (CES), Lenovo Tech World, Austrian Parliament and Zero Project Conference. ALS is a disease that affects the nerve cells that control your movement, leading to increasing muscle weakness and eventual paralysis.There is no cure, but treatments are available to manage symptoms and help improve quality of life. |

# 2 Use Case Description

## 2.1 Description

Imagine a world where technology bridges the gap between silence and expression, giving everyone a voice. Today, over 500 million people worldwide face barriers to communication, a fundamental human right that shouldn't be locked behind expensive medical devices.

The Scott-Morgan Foundation (SMF), in collaboration with leading technology partners, is revolutionizing assistive technology through an ecosystem-based approach that restores not just speech, but human identity. By designing for extreme cases – people with ALS who may retain only eye movement control – we're creating solutions that serve the entire spectrum of communication needs. Like curb cuts designed for wheelchairs that benefit everyone with wheels, our solutions for the most severe disabilities create technology advances that improve accessibility for all.

Our approach brings together eye-tracking (IRISBOND), advanced LLMs for prediction, personalized voice synthesis (ElevenLabs), and avatar technology (D-ID). Lenovo provides the hardware platform and interface development, while SMF serves as the catalyst, bringing these technologies together in a unified ecosystem. Rather than creating yet another siloed solution, we're establishing open standards and connection points that allow diverse assistive technologies to work together seamlessly.

Our solution is designed to address critical communication barriers:

Current Communication Landscape:

- Over 500 million people worldwide face communication barriers
- Only 18.2% of students needing communication support receive Augmentative and Alternative Communication (AAC)
- Merely 4.8% access speech-generating devices

Technological Innovation:

- 80% cost reduction compared to existing medical-grade solutions
- Open ecosystem approach connecting previously siloed technologies
- Adaptive input systems with advanced machine learning predictions

Scalability Approach: Our platform is designed with modularity and adaptability at its core:

- Personalized voice synthesis from limited samples

- Eye-tracking technology compatible with multiple platforms
- Adaptive interfaces supporting diverse user needs
- Open-source integration framework to enable continuous improvement

Collaborative Development:

- Partnerships with leading technology companies (Lenovo, ElevenLabs, D-ID, IRISBOND and many others)
- Testing and validation with ALS users and disability organizations
- Commitment to creating an inclusive, accessible communication solution

The platform emphasizes human-centric design, accessibility, and dignity, with on-device processing for privacy and real-time performance. We've successfully demonstrated this solution in high-profile venues:

- At Lenovo Tech World: Erin Taylor (ALS) introduced Lenovo's CEO using the technology
- At Austrian Parliament: Our avatar system enabled communication in Viennese German and gave an ex-parliamentarian with paralysis a voice
- With ALS families: A father reading stories and singing lullabies to his children

## ITU Standard Alignment

Our solution directly aligns with critical ITU telecommunications standards:

- ITU-T F.930 (06/2018): Guidelines for developing accessible telecommunication services
- ITU-T F.931 (09/2019): Telecommunications accessibility checklist for people with disabilities
- ITU-T Y.4116/L.1630 (09/2020): Requirements for AI-enabled assistive technologies

Research confirms the concerning gap in assistive communication technology access. According to the American Speech-Language-Hearing Association, despite the critical need for communication support, less than 20% of U.S. students who need communication assistance receive any form of AAC. This accessibility gap widens further when considering the millions more who remain undiagnosed or without proper assessment, particularly in underserved communities.

Use Case Status: Successfully demonstrated with working technology in real-world settings; part of ongoing development initiative.

**Partners**

The Scott-Morgan Foundation has established a revolutionary collaboration model uniting technology leaders through strategic social impact partnerships:

**Lenovo**: Provides hardware, interface design, and global reach. Developed the circular keyboard interface, provided testing hardware, led prototype integration, and showcased the technology at Tech World.

**ElevenLabs**: Contributed voice synthesis technology creating personalized voices from limited samples. Committed to distributing one million synthetic voices for people with disabilities (valued at $99 million).

**D-ID**: Provides avatar technology preserving visual identity and expression, enabling lifelike representation from limited video or photos.

I**RISBOND**: Contributes eye-tracking technology through their Hiru camera, providing precise tracking across platforms and conditions.

**NVIDIA**: All of our partner companies are either founding members of NVIDIA's Inception program (D-ID and ElevenLabs), or we have migrated them into the program. Our CTO, Thorsten Stremlau, formerly Lenovo's CTO and part of Stephen Hawking's IT team before working with our founder Dr. Peter Scott-Morgan, is now at NVIDIA. Working with the foundation's Chief Engineer and Distinguished Engineer Eric Kern at Lenovo, the two provide AI expertise and enable real-time processing of complex AI models across our ecosystem of technologies.

Zero Project, CES, Lenovo Tech World: Provide platforms for global demonstrations, including at Austrian Parliament and UN headquarters in Vienna.

Central to our success are multiple ALS participants serving as both testers and ambassadors, ensuring solutions address real-world needs.

## 2.2 Benefits of the Use Case

The AI-powered communication platform offers significant benefits by improving mental health, quality of life, and healthcare outcomes through enhanced communication capabilities. It enables individuals to effectively communicate symptoms and needs to caregivers and medical professionals, ensuring better healthcare support. In the realm of education, the platform provides essential communication tools that facilitate learning and participation, supporting multiple languages and adaptive interfaces to ensure inclusive education for all. The platform showcases innovative infrastructure through the open ecosystem integration of cutting-edge AI, creating a modular architecture that can continuously adapt to emerging technologies. By making advanced communication tools accessible regardless of economic status and reducing costs by 80% compared to medical-grade solutions, it helps bridge economic disparities. Furthermore, the platform exemplifies successful multi-stakeholder collaboration across nonprofits, corporations, startups, and users, laying the foundation for ongoing cross-sector development and fostering a collaborative environment for future advancements.

## 2.3 Future Work

Future work will focus on three critical areas:

Open Standards and Platform Development:

- Develop open-sourced platform and standards enabling assistive technologies to plug and play
- Create a framework allowing diverse technologies to work together seamlessly
- Implement testing with people who have severe disabilities
- Optimize for device independence and on-device processing

Reference Implementation and Standards:

- Develop documentation and reference implementations
- Create standardized APIs for diverse input devices and AI models

- Build demonstration environments showing integrated technologies
- Contribute to ITU standards for AI-powered assistive technologies

Community Support and Testing:

- Establish testing programs with ALS clinics and disability organizations
- Build a developer community around the open-source platform
- Create resources for ongoing maintenance and support

Resources would support engineering, infrastructure, testing equipment, and community programs ensuring continued evolution and accessibility.

## 3 Use Case Requirements

- **REQ-01**: Multimodal Input Support

  Support multiple input methods, primarily eye-tracking for severe mobility impairments, with adjustable sensitivity for different capabilities.

- **REQ-02**: Personalized Voice Synthesis

  Generate natural-sounding voices from limited samples, maintaining emotional expressiveness and personal speech patterns.

- **REQ-03**: Visual Identity Representation

  Provide realistic avatar generation preserving identity and expressions, enabling non-verbal communication cues.

- **REQ-04**: Low Latency Response

  Enable real-time communication (<500ms) from eye movement to voice/avatar output.

- **REQ-05**: Privacy Preservation

  Process personal data on-device when possible, with usable voice and avatar models without continuous cloud connectivity.

- **REQ-06**: Accessibility Across Economic Barriers

  Deploy on consumer hardware to reduce costs by 80% compared to specialized medical devices.

- **REQ-07**: Open Standards and Interoperability

  Support open standards for easy integration of new technologies and interoperability with existing systems.

- **REQ-08**: Cross-Platform Support

  Function across multiple platforms, including Windows computers and iPad tablets.

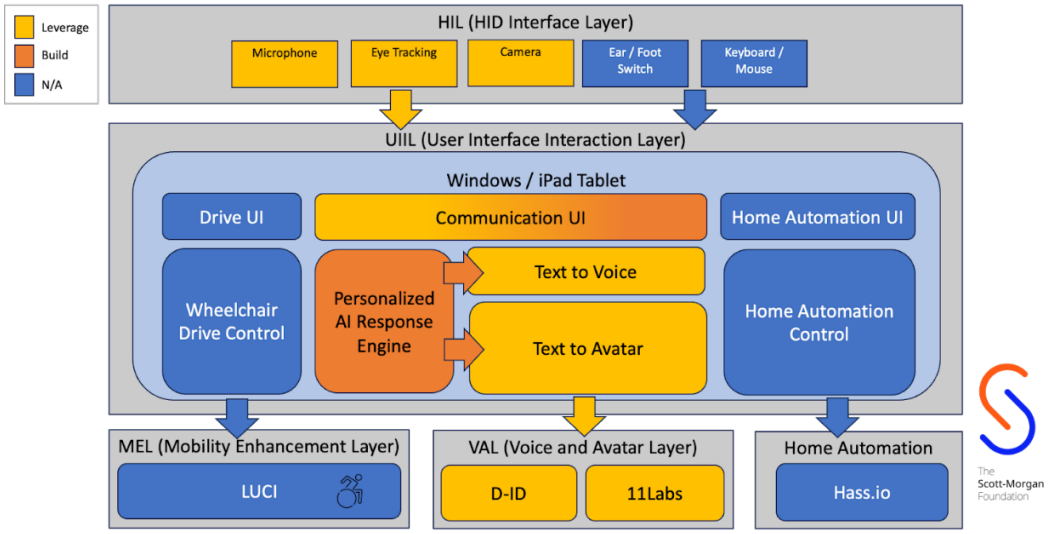
- **REQ-09**: Predictive Communication

  Incorporate AI-powered prediction to accelerate communication, reducing selections needed for common phrases.

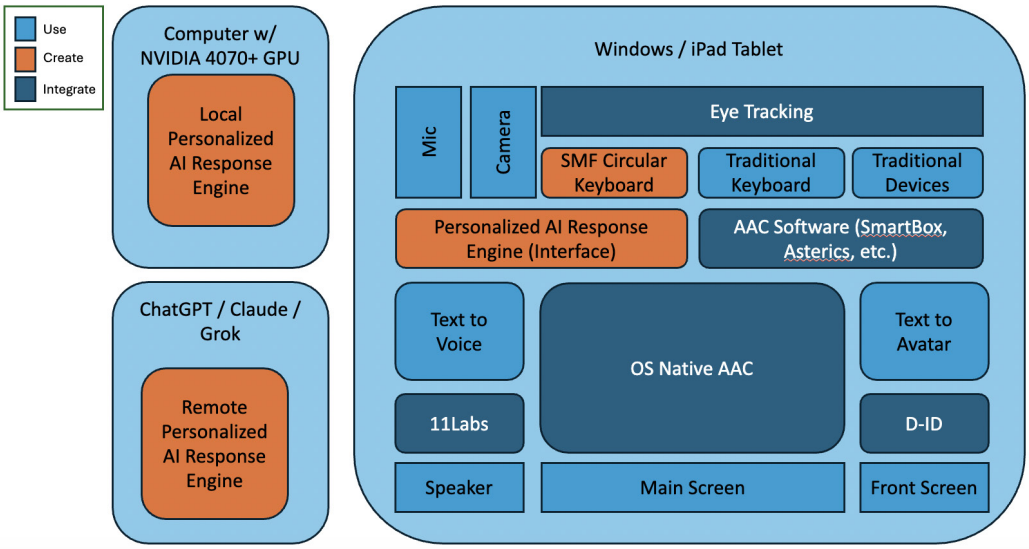
- **REQ-10**: Modularity

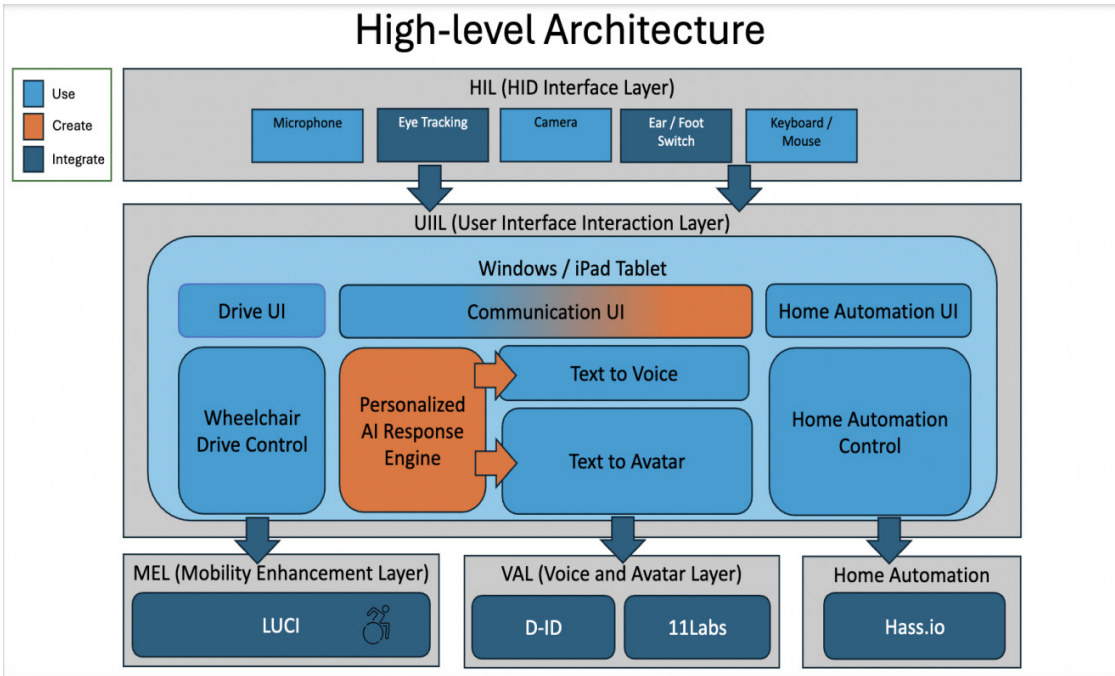
  Enable customization based on individual needs and to accommodate progressive disability conditions.

# 4 Sequence Diagram




## Detailed Architecture






The architecture diagrams illustrate our vision for integrated assistive technology (see below in references):

1) HIL (HID Interface Layer): Handles input methods (eye tracking, microphone, camera, switches, keyboard/mouse)
2) UIIL (User Interface Interaction Layer): Provides user interfaces (Communication UI, Drive UI, Home Automation UI, Text-to-Voice/Avatar, AI Response Engine)
3) Specialized Layers: Handle specific functions (Mobility Enhancement, Voice and Avatar Layer, Home Automation)

Rather than building everything from scratch, we're creating an open ecosystem where technologies can "plug and play," freeing users from proprietary all-in-one solutions.

## 5 References

[1] World Health Organization. (2024). Assistive technology fact sheet. Available at: https://www.who.int/news-room/fact-sheets/detail/assistive-technology

[2] American Speech-Language-Hearing Association. (2024). Augmentative and Alternative Communication (AAC). Available at: https://www.asha.org/practice-portal/professional-issues/augmentative-and-alternative-communication/

[3] Scott-Morgan Foundation. (2024). Official website. Available at: https://www.scottmorganfoundation.org/

[4] Scott-Morgan Foundation. (2024). Scott-Morgan Foundation Unveils Revolutionary AI-Powered Assistive Technology in Collaboration with Lenovo at Tech World '24. Press release.

[5] Lenovo. (2025). Transforming the frontiers of assistive tech with visionary AI at Zero Project Conference 2025. Lenovo StoryHub.

[6] Lenovo. (2024). AI-powered communication solution for people living with ALS revealed at Lenovo Tech World. Available at: https://news.lenovo.com/pressroom/press-releases/ai-powered-communication-solution-for-people-with-als-revealed-tech-world/

[7] ElevenLabs. (2024). Impact Program Announcement. Available at: https://elevenlabs.io/blog/impact-program-announcement

[8] International Telecommunication Union. (2018). ITU-T F.930: Guidelines for developing accessible telecommunication services. Geneva: ITU.

[9] International Telecommunication Union. (2019). ITU-T F.931: Telecommunications accessibility checklist for people with disabilities. Geneva: ITU.

[10] International Telecommunication Union. (2020). ITU-T Y.4116/L.1630: Requirements for AI-enabled assistive technologies. Geneva: ITU.

# Use Case 16: AI Cyber Security Edge and Cloud based Intrusion Detection System for IoT

**Organization**: Addis Ababa Science and Technology University

**Country**: Ethiopia

**Contact person**:

Mr. Kibremoges Fenta , email: kibremogesf@gmail.com

## 1 Use Case Summary Table

| Item | Cyber Security |
|---|---|
| Category | Productivity |
| Problem Addressed | The increasing vulnerability of interconnected IoT devices due to weak access point security and the lack of real-time, intelligent monitoring mechanisms. |
| Key Aspects of Solution | • AI-powered real-time intrusion detection deployed at the edge, like routers<br>• Cloud-base analytics for flagged threats<br>• Continuous training and fine tuning for accuracy and speed<br>• Subscription/contact based access control for traffic monitoring<br>• Focus on privacy, safety, and digital trust |
| Technology Keywords | Edge AI, Intrusion Detection, IoT Security,<br>Cyber Threat Intelligence, Real-time Monitoring, Deep Learning, CNN-LSTM |
| Data Availability | Modified and publicly available on kaggle Link https://doi.org/10.34740/kaggle/dsv/11932656 |
| Metadata (Type of data) | • Tabular data (CSV format)<br>• Multi-class network traffic( benign + multiple attack types) |
| | • Models trained on class-balanced and cleaned version of CICIoT2023<br>• CNN-LSTM and Transformer model used for classification<br>• Evaluating using standard train-test split<br>• Accuracy, F1-score, recall, and inference speed optimized |

(continued)

| Item | Cyber Security |
|---|---|
| | • Simulated deployment at Edge for real-time inference<br>• Traffic sink to cloud module for post event analysis<br>• Planed lab-scale implementation in smart home/test IoT network |
| Code Repository | GitHub, Coming soon |

# 2 Use Case Description

## 2.1 Description

*Ensuring safety and privacy in wearable* health devices, medical equipment, and remote monitoring systems has a high impact on.

Globally the use of digital systems are increasing daily so to make this technology advancement usage ethically well and as intended we need to come up with a general solution for digital Cyber security intrusion. It's a new approach to innovate on this rapidly growing sector and cutting edge technology that makes it align with SDG 9.

Also it decreases in equality that most of the threat comes from absences or level of digital infrastructure knowledge this means it can solve this problem by giving all equal protection regardless of their expertise level and knowledge.

We can make cities sustain well if we build all the digital infrastructure trust worthiness higher, and the community as a whole can use and exercise it with confidence.

In climate change most of our data gathered through some sort of sensors on the network, this needs an urgent action to make sure all the data are reliable and not corrupted by someone, to predict and work better.

Securing the network environment maintains peace and trust to all rely on strong institutions that lead to it.

## 2.2 Benefits of the use case

Currently on a global scale, the use of digital systems is increasing daily. So, to make this technological advancement go *{please fill* well with our best interest, we need to come up with general ways of interaction *these}* with the systems in general and specific activities that may affect our health and well-being.

in the education sector, we can have good education by interacting with a computer system freely and in an easier way. This leads to innovating new approaches to solve problems we face daily.

Nowadays, the use of such a digital system is very crucial on a global scale. For users with disabilities using such systems not only put them on equal grounds to the digital platform but also develop and share their natural gifts as a person - as the saying goes and we believe it's true - they are differently enabled. I believe our proposed system will solve such an urgent problem in the community and address it well.

## 2.3 Future Work

1. Data Collection, to make it more general and work for most digital devices we may need gathering more data not only in amount but also from different devices and threat types.
2. Model Development, this would enable us to build and fine tune a more general and novel detection system.
3. Model development and training.

With our model we are trying to detect and identify the threat type as well

# 3 Use Case Requirements

**REQ-01: Real-time Edge Monitoring**

It is critical that the system performs real-time traffic inspection at the edge level (like routers, IoT gateways) to immediately detect suspicious activity.

**REQ-02: Cloud-Based Analytics and Response**

It is critical that flagged or suspicious traffic is automatically forwarded to the cloud for deep analysis, logging, and correlation with known threat intelligence databases.

**REQ-03: Accuracy and Continuous Learning**

It is expected that the system maintains high accuracy and low false positives by continuously retraining its AI models using updated, diverse datasets.

**REQ-04: Privacy-Aware Traffic Monitoring**

It is critical that the system ensures data privacy and user consent, using a subscription-based or contract-based model to authorize traffic inspection.

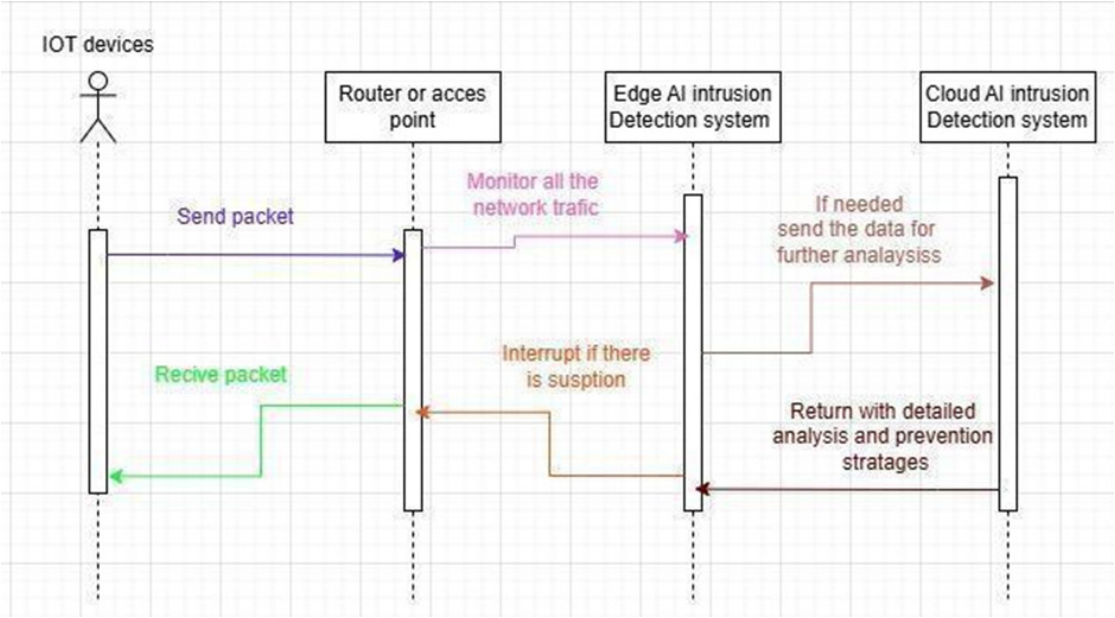
**REQ-05: Scalability and Deployment Flexibility**

It is expected that the solution can be deployed in various environments (e.g., smart homes, offices, healthcare, manufacturing) with minimal configuration overhead.

**REQ-06: Performance and Inference Speed**

It is critical that the IDS provides fast inference times with minimal latency, allowing seamless operation without impacting network performance.

## 4 Sequence Diagram




## 5 References

[1] Hakan Can Altunay, Zafer Albayrak, A hybrid CNN+LSTM-based intrusion detection system for industrial IoT networks, Engineering Science and Technology, an International Journal,Volume 38, 2023, 101322, ISSN 2215-0986, https://doi.org/10.1016/j.jestch.2022.101322

[2] Uwadia Anthony. O. "CNN-based Network Intrusion Detection and Classification Model for Cyber-Attacks." Volume. 9 Issue.7, July - 2024 International Journal of Innovative Science and Research Technology (IJISRT), www.ijisrt.com. ISSN - 2456-2165, PP:- 1839-1847, https://doi.org/10.38124/ijisrt/IJISRT24JUL1158

# Use Case 17: The future of Work and Attached Intelligence for Knowledge Workers: A Framework for Cognitive Resilience and Mindful Productivity

**Organization**: Etcera by Akimilabs

**Country**: Canada

**Contact Person(s)**:

David Agahchen, david.agahchen@akimilabs.com

Anissa Agahchen, anissa.agahchen@akimilabs.com

## 1 Use Case Summary Table

| Item | Details |
|---|---|
| Category | Mindful productivity, cognitive resilience, human-centered decision support. |
| Problem Addressed | Knowledge workers face growing decision fatigue and fragmented workflows due to high cognitive load, constant context switching, and reactive task cycles—reducing their capacity for deep work, alignment, and long-term resilience. |
| Key Aspects of Solution | – Think-Do-Reflect productivity framework.<br>– Attached Intelligence Multi-agent System (AIM-1).<br>– Context-aware support and reflection prompts.<br>– Modular orchestration of specialized AI agents.<br>– Goal alignment, emotional modeling, and knowledge graph integration. |
| Technology Keywords | Attached Intelligence, Large Language Model (LLM) -powered AI, cognitive augmentation, natural language understanding, AI-assisted reflection, mindful productivity, multi-agent orchestration, context-aware memory. |
| Data Availability | Private (collected and stored securely in Etcera platform; not publicly shared.) |

(continued)

| Item | Details |
|---|---|
| Metadata (Type of Data) | Text (thoughts, reflections, actions), temporal context, emotional states, intentions, decision links, topic threads. |
| Model Training and Fine-Tuning | Retrieval-augmented generation and memory-enhanced personalization. No model retraining required. AIM-1 is a modular system of coordinated AI agents uses LLM inference combined with user feedback, contextual memory, and modular agent coordination. |
| Testbeds or Pilot Deployments | Internal MVP in testing at with knowledge professionals across product, operations, development, design, and other cognitively demanding roles—supporting clarity, resilience, and access to mindful productivity in modern work. Available at [8]. |
| Code repositories | Private (in development.) |

# 2 Use Case Description

## 2.1 Description

This use case introduces *Attached Intelligence* – a new paradigm for AI that works with the user, not for or against them. Unlike traditional automation or command-based agents, Attached Intelligence is collaborative: it supports users through structured thinking, intentional action, and guided reflection [2], while preserving human agency and adaptability.

Modern professionals and knowledge workers are inundated with decisions every day—ranging from routine to highly strategic. While some are simple, they still require attention and mental energy [1]. Others demand deep context, domain knowledge, and the ability to weigh competing priorities. These decisions are rarely made in isolation; they are often layered with cross-functional considerations, time pressure, and shifting goals.

Even routine decisions—like prioritizing tasks, responding to messages, or switching between tools—consume the limited cognitive bandwidth available. This reduces the attention and focus professionals can dedicate to deeper, high-impact work and strategic decision-making [1]. The result is a fragmented workday, rising decision fatigue, and a sense of being busy without meaningful progress.

Our aim is to support *workers in knowledge labor sectors who are vulnerable to burnout and disruption* in the evolving AI-enhanced workplace. By enhancing cognitive clarity and reflective capacity, we help them make more intentional, high-quality decisions. Rather than just managing tasks, we support the entire cognitive loop of thinking, doing, and reflecting [1] [2] so that decisions become more aligned, agile, and effective over time.

We are developing a productivity framework embedded in a Think-Do-Reflect workflow. It is powered by the Attached Intelligence Multi-agent System (AIM-1) – a modular, LLM-powered system designed to operate as a collaborative thinking partner. AIM-1 enables context-aware,

reflective decision support by orchestrating specialized agents that work together to guide users through structured thinking, actionable planning, and ongoing reflection—supporting user agency at each step without taking over the decision.

This attached system integrates into the user's flow without taking over or overwhelming them. The assistant, as part of AIM-1, helps users structure their thinking, clarify actions, and reflect on outcomes. Conversations and actions are continuously broken down into evolving topics and stored in a personalized, dynamic knowledge graph. This allows users to zoom in or out of their context—whether they need a high-level overview or a deep dive into a specific topic or decision. As users provide feedback, AIM-1 refines its support through context-based memory and retrieval-enhanced learning, becoming more accurate and helpful over time.

Expected Impacts and Benefits:

- Reduced decision fatigue by streamlining workflows and offloading routine choices.
- Increased clarity and confidence in both daily and high-stakes decisions.
- More time and focus for deep work and strategic thinking.
- Continuous personal and professional growth through guided reflection.
- Zoomable context views, from strategic themes to granular decisions.
- Context-aware AI assistance that adapts to the user and aligns with human-centered values.
- Emotionally-aware guidance that supports transitions from reactive to intentional behavior.
- A scalable model for mindful productivity and decision-making in knowledge work.

Use Case Status:

The Etcera MVP is currently in internal testing with selected knowledge professionals across product, development, operations, and creative roles. We are actively accepting early access requests from individuals and organizations working in decision-intensive environments, with priority given to those aligned with SDG 8 objectives. Early feedback is being used to refine contextual memory, agent coordination, and reflection guidance features ahead of broader pilot deployment.

## 2.2 Benefits of the Use Case

Etcera's Attached Intelligence Multi-agent System (AIM-1) enhances human decision-making, promotes inclusivity, and fosters transparent institutions through responsible AI assistance. Benefits include:

- Improved economic productivity through reduced cognitive overload and streamlined decision-making. This may contribute to higher economic output per worker [3][7].
- Support for full and productive work through improved mental energy and clarity, helping professionals stay focused and resilient.
- Greater accessibility to cognitive support tools via natural-language interfaces, which can help reduce barriers for non-technical users and support equitable access [6].
- Fostered trust in AI-assisted work environments by emphasizing traceable decision support and explainability, aligning with transparent governance practices [6].

- Support for inclusive, participatory decision-making through guided reflection and user agency.through reflection-driven, human-centered design [6].

## 2.3 Future Work

Etcera envisions an Attached Intelligence Era—a future where AI enhances human agency, builds cognitive resilience, and empowers individuals and communities to thrive. This vision will be realized through both technical innovation and human-centered partnerships grounded in sustainable development.

Our immediate goal is to advance the submitted use case—a productivity framework for knowledge workers powered by the AIM-1 system. We aim to refine the system through testing, research collaborations, and sustainable scaling.

### Invitation-Based Pilot for Individuals

The next phase involves launching an invitation-only pilot program for individuals engaged in decision-intensive work. This pilot will test how AIM-1 supports cognitive clarity, decision quality, emotional resilience, and reflective practice in real-world settings.

### Measuring Mindful Productivity

We are developing a measurement framework to evaluate how the Think-Do-Reflect cycle impacts user well-being and performance. This framework will help us quantify cognitive productivity and emotional resilience in meaningful, actionable ways.

### System Enhancements

Ongoing development of the AIM-1 system will focus on improving decision support, expanding contextual memory, and supporting diverse input methods—all guided by real-world feedback and aligned with global standards for responsible AI.

### Research Partnerships

We are seeking partnerships with researchers and scholars to co-develop new metrics, training approaches, and best practices that support reflective decision-making and sustainable knowledge work.

# 3 Use Case Requirements

Technical Requirements

- **REQ-01**: The system must support multi-modal capture (e.g., text, voice, notes) to record thoughts, actions, and reflections in real time.
- **REQ-02**: The platform must integrate a modular, multi-agent architecture (AIM-1) to orchestrate personalized Think Do Reflect cycles.
- **REQ-03**: The assistant must provide real-time natural language processing and retrieval-augmented responses to surface context-relevant insights.
- **REQ-04**: The system must ensure explainable, traceable reasoning behind AI suggestions to support user trust and decision transparency.

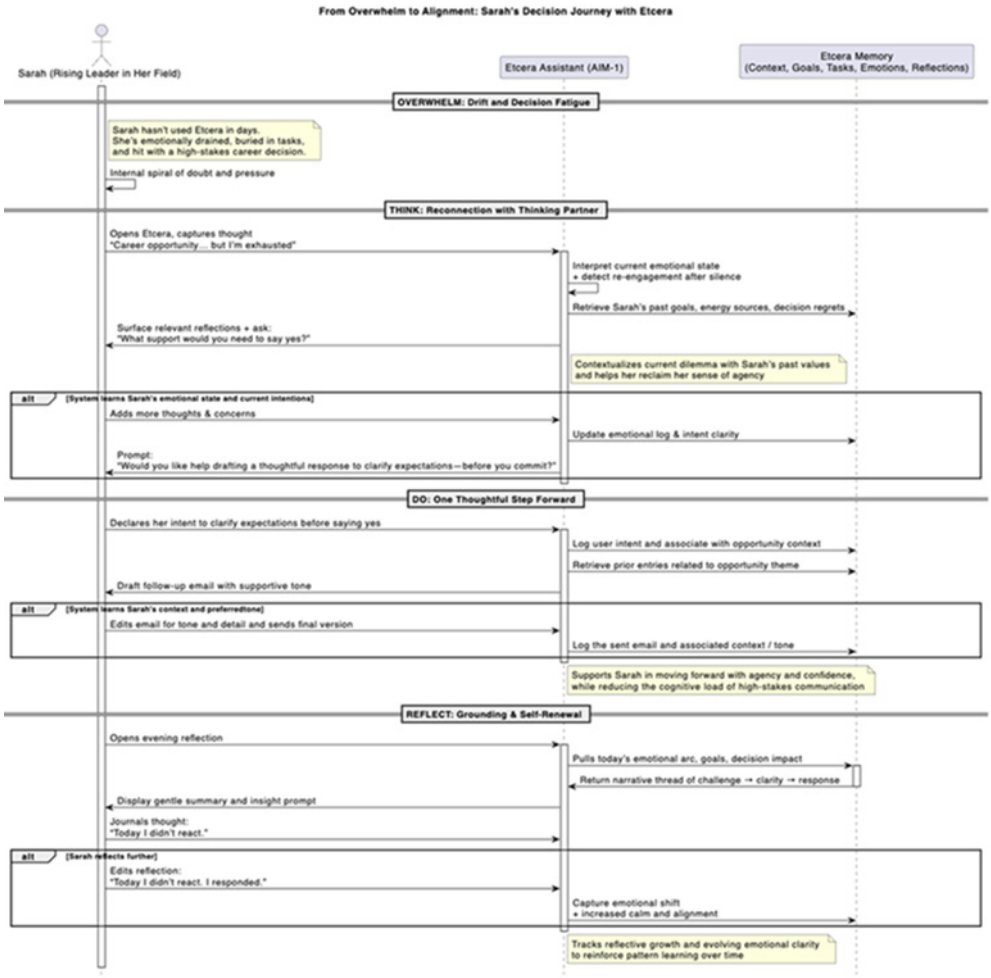
- **REQ-05**: User entries must be organized by topics, intentions, and thought bundles to enable meaningful structure and recall.

Operational & Functional Requirements

- **REQ-06**: The assistant must operate across platforms (web, desktop, mobile), ensuring continuity across locations and work contexts.
- **REQ-07**: The platform must support visual goal alignment and progress mapping to help users prioritize what matters most.
- **REQ-08**: The system must encourage structured self-reflection through smart prompts, micro-win recognition, and habit tracking.
- **REQ-09**: The AI must continuously personalize its guidance based on evolving user context and feedback without requiring manual retraining.
- **REQ-10**: The system must track energy, attention, and emotional patterns over time to help users build cognitive resilience and avoid burnout.

## 4 Sequence Diagram

This scenario follows Sarah, a rising leader facing a high-stakes decision moment and showcases how the system scaffolds clarity, agency, and emotional alignment.




## 5 References

[1] Use case organization official website: https://etcera.com

[2] Zimmerman, B. J. (2000). *Attaining self-regulation: A social cognitive perspective*. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), Handbook of self-regulation (pp. 13–39). San Diego, CA: Academic Press.

[3] Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: W. H. Freeman.

[4] Autor, D. H. (2015). *Why are there still so many jobs? The history and future of workplace automation*. Journal of Economic Perspectives, 29(3), 3–30.

[5] Bessen, J. E. (2018). *AI and jobs: The role of demand*. NBER Working Paper No. 24235.

[6] Saetra, H. S. (2023). *AI for the Sustainable Development Goals*.

[7] International Labour Organization. (2023). *Generative AI and jobs: A global analysis of potential effects on job quantity and quality*.

[8] OECD. (2024). *OECD AI Principles*. Organisation for Economic Co-operation and Development.

# Use Case 18: AI-Powered Talent Matching Platform

**Organization**: Dubai Government Human Resources Department Director General Office

**Country**: United Arab Emirates

**Contact person(s)**:

Aamena Al Shamsi - **aalshamsi@dghr.gov.ae**

**Eman Al Ali -** ealali@dghr.gov.ae

**Soha Jawabreh -** sjawabreh@dghr.gov.ae

## 1 Use Case Summary Table

| Category | Productivity |
|---|---|
| Problem Addressed | this use case addresses the challenges of workforce allocation by accurately matching CVs with job descriptions. |
| Key Aspects of Solution | An AI-powered semantic search engine designed to address the key challenges of workforce allocation by accurately matching CVs with job descriptions through Running the AI system on the provided CVs to generate insights and match job opportunities. |
| Technology Keyword | search engine, workforce allocation |
| Data Availability | private |
| Meta Data | text based bio data. |
| Model Training and Fine-tuning | N/A |
| Testbeds of Pilot Deployment | N/A |
| Code Repositories | N/A |

## 2 Use Case Description

### 2.1 Description

An AI-powered semantic search engine designed to address the key challenges of workforce allocation by accurately matching CVs with job descriptions through Running the AI system on the provided CVs to generate insights and match job opportunities.

#### Partner

TRANSCEND TECHNOLOGIES L.L.C

## 2.2 Benefits of the use case

AI optimizes workforce allocation, reduces unemployment, and bridges skills gaps for sustainable economic growth.

AI accelerates industrial innovation by connecting skilled professionals to high-demand industries and infrastructure projects.

AI-driven, bias-free hiring ensures diverse, inclusive employment opportunities, reducing socio-economic and gender disparities.

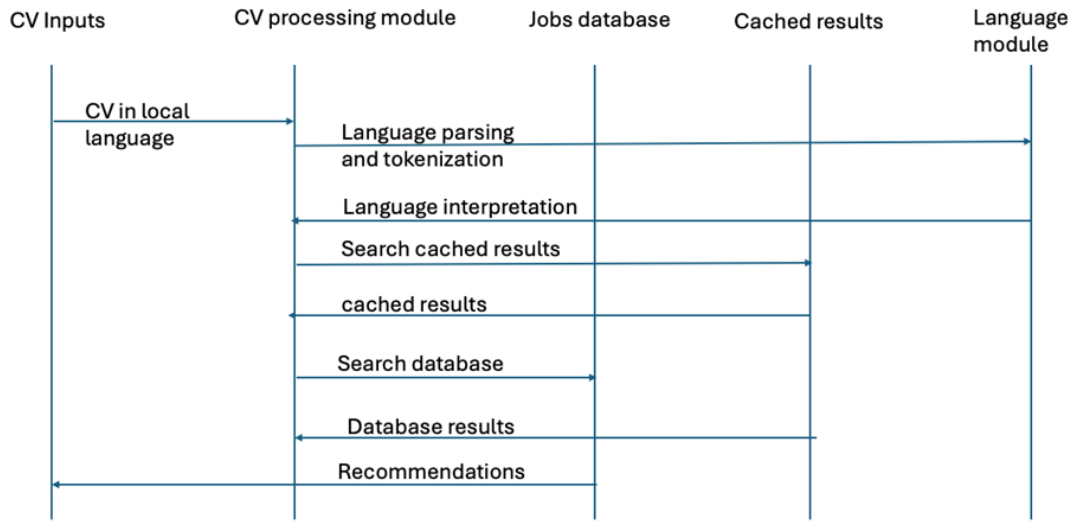
## 2.3 Future Work

Future works include Model development, Model Training, Data analysis & studies.

# 3 Use Case Requirements

- it is critical that the AI system reads and converts CV to structured format
- it is critical that the solution includes semantic search to find related jobs for CVs.
- it is expected that the solution outputs reranked priorities.
- it is expected that the solution implements Large Language Models (LLMs) for scoring and rewriting CVs.
- it is expected that the solution supports Arabic and English and reuses cached results for optimal speed.

# 4 Sequence Diagram




# 5 References

# 4.5. Category 5: Manufacturing

## Use Case 1: Morphing-BOT: Revolutionizing Global Infrastructure with AI-Powered Smart Inspection and Predictive Maintenance

**Organization**: Morphing I

**Country**: South Korea

**Contact Person(s)**:

Kevin Kwon , dohoon57@morphingi.com

Kiyoung Kim, kiyoungkim@morphingi.

Kyeongmin Lee, km0107@morphingi.

### 1 Use Case Summary Table

| Item | Details |
|---|---|
| Category | Manufacturing |
| Problem Addressed | Aging and buried water pipelines are difficult to access and inspect, especially in complex, curved, or narrow sections. Traditional inspection methods are labor-intensive, invasive, and limited in real-time data acquisition. This results in undetected anomalies, delayed maintenance, and increased risk of leaks, bursts, and water losses in urban infrastructure. |
| Key Aspects of Solution | Morphing I, provides an AI-powered pipeline inspection robot, "Morphing-BOT," integrated with IoT sensors and a digital twin platform.<br>The robot navigates narrow and curved pipes autonomously, detects anomalies, collects multi-modal data (acoustic, vibration, image), and transmits real-time insights to a digital twin system for predictive maintenance. |
| Technology Keywords | AI Robot, Smart Infrastructure, Digital Twin, Predictive Maintenance, Pipe Inspection, Acoustic AI, Visual AI |
| Data Availability | Private (Data collected from live testbeds in South Korea. Over 110,000 labelled data points collected and validated by TTA[1]) |
| Metadata (Type of Data) | Image, Acoustic (hydraulic noise), Lidar, Environmental (pressure, humidity) |
| Model Training and Fine-Tuning | Supervised learning for anomaly classification using CNNs and transformers; domain-specific fine-tuning for different pipeline types and countries; on-device lightweight models for real-time inference and off-device training for global optimization. |

(continued)

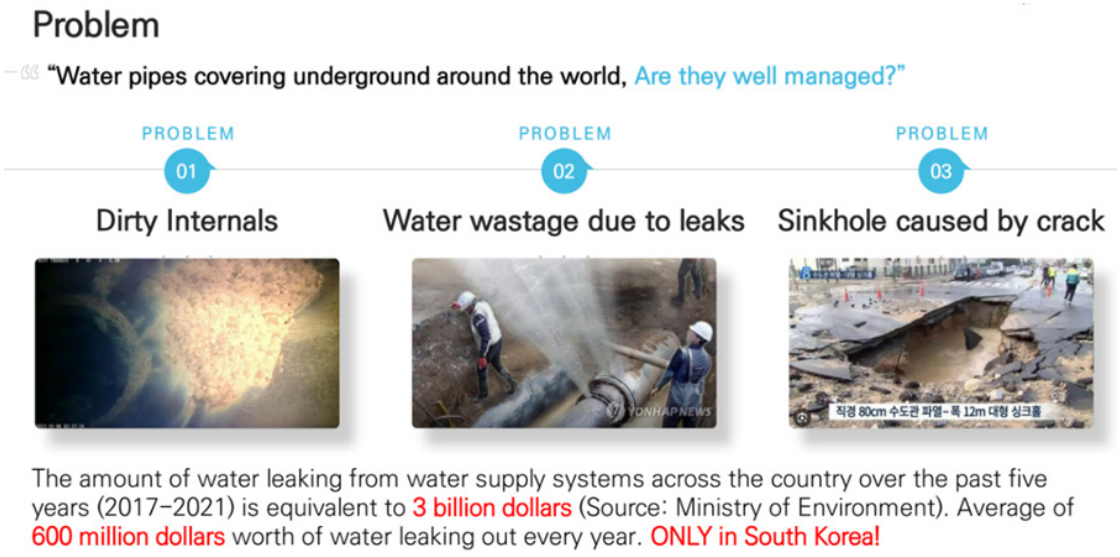
| Item | Details |
|---|---|
| Testbeds or Pilot Deployments | 30km live testing with municipalities in Jeju, Jindo, Jeonju, and Siheung<br>PoC (Proof of Concept) requests in Saudi Arabia (200km), Canada (AQUADATA), and Qatar (Digital Enterprises)<br>NIPA UAE partnership and KIC Silicon Valley entry confirmed |

# 2 Use Case Description

## 2.1 Description

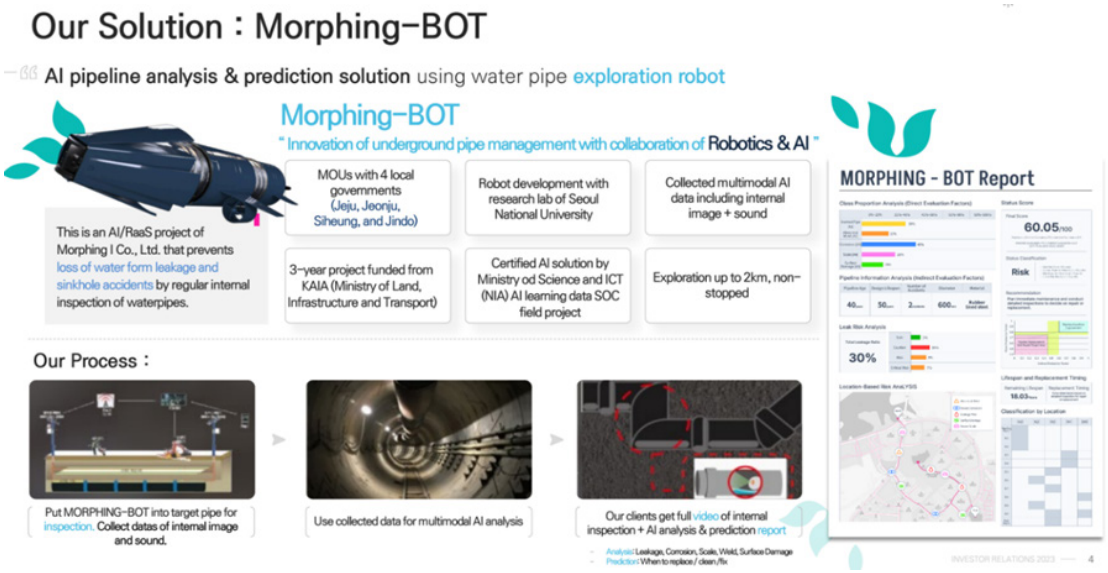
### Problem

Aging urban infrastructure, particularly underground water pipelines, often suffers from undetected leaks, cracks, and failures. Traditional inspection methods are labor-intensive, slow, and lack real-time data integration, leading to inefficient maintenance and costly repairs.




### Objective

The objective to develop and deploy an AI-powered robotic system combined with IoT sensors and digital twin technology to enable real-time asset monitoring, early detection of anomalies, and predictive maintenance for water pipelines.

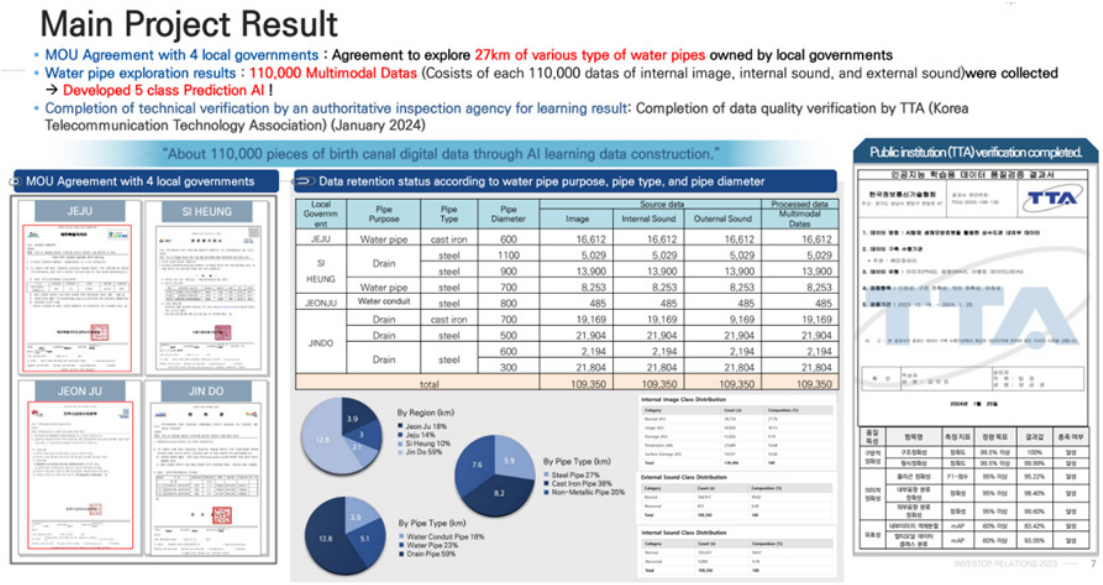



## Technological approach

The solution integrates an AI exploration robot equipped with IoT multi-sensors for acoustic, vibration, and structural data collection. This data is fused into a digital twin platform that offers real-time visualization, anomaly prediction via machine learning algorithms, and automated asset management functionalities, providing a holistic maintenance solution [2].

## Use Case Status

- Completed successful pilot projects with multiple municipalities and public infrastructure operators in South Korea.
- Validated data quality through third-party certification (TTA) and real-world pipeline inspections covering over 30 kilometers.
- Secured LOIs and partnerships with global clients in Saudi Arabia, Qatar, UAE, Canada, and the United States.
- Currently transitioning into early commercialization with PoC projects underway in North America and the Middle East.
- Preparing for scale-up with overseas subsidiaries established in the United States and Southeast Asia, and plans for GCC market entry through UAE and Saudi offices.

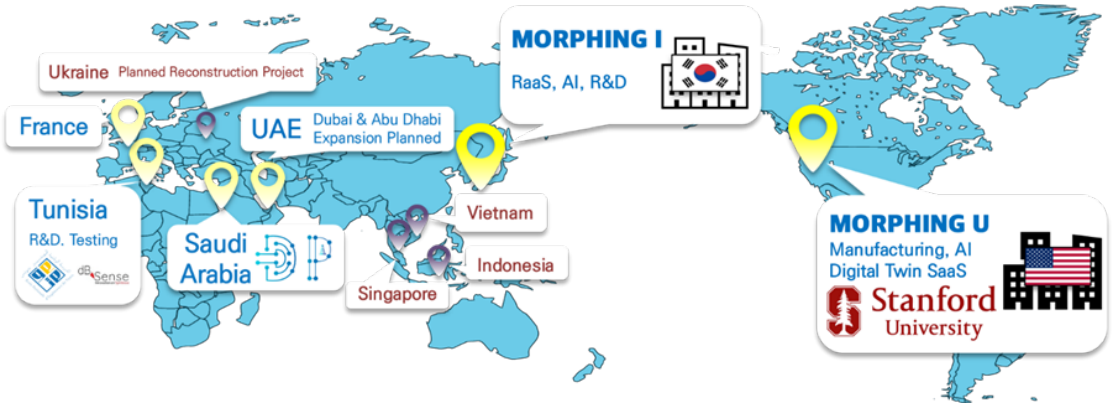



## Project Piloting and Scaling-up for Global Commercialization

Our project has successfully completed multiple pilot deployments across South Korea and is now entering a global expansion phase with several PoCs and international research collaborations underway:

- **Singapore**: PoC planning is in progress with **DSTA (Defence Science and Technology Agency)** and local authorities for deployment in urban pipeline systems.
- **Canada**: A joint PoC is scheduled with **AQUADATA** to demonstrate our AI-powered inspection robot on North American pipelines.
- **Saudi Arabia**: A local partner has requested a formal quotation for a **200-kilometer pipeline inspection**, with field deployment now under preparation.
- **United States**: Strong interest was received during our participation in the **SelectUSA Investment Summit**, including:
- Direct meetings with the governors and public officials of **Montana** and **Richmond, Virginia**, with follow-up visits and potential municipal integration underway.
- A joint PoC and research collaboration with **Stanford University** will be launched to validate and localize our solution for U.S. infrastructure. (have LOI and MOU)
- **MENA Region**: An MoU and LOI have been signed with **ENIT (National Engineering School of Tunis)** to jointly conduct field PoC trials and R&D in North Africa and the Middle East.

These activities mark our transition from validated pilot projects to real-world implementations and international commercialization through partnerships, PoCs, and collaborative research.




## Operational Mode

Morphing-BOT operates in semi-autonomous mode, autonomously navigating and mapping pipelines using sensor fusion (IMU, Lidar, acoustic sensors). Remote control is available for manual intervention in complex or highly obstructed areas.

## Data Utilization

The datasets are utilized to train AI models for anomaly detection (cracks, leaks, deformation), predictive maintenance prediction, asset condition rating, and navigation optimization (route mapping and obstacle avoidance).

## AI Architecture & On-Device Functions

Our intelligent inspection robot uses multimodal sensor data, including high-resolution images, vibration signals, acoustic measurements, and IMU data, to detect pipeline anomalies. To process this data, we employ Convolutional Neural Networks (CNNs) for image-based defect detection, Recurrent Neural Networks (RNNs) for time-series vibration and acoustic analysis, and Transformer-based architectures for multimodal feature fusion and anomaly prediction. The on-device AI facilitates real-time processing and decision-making, performing immediate obstacle detection and avoidance, defect recognition and classification without external computing support, route adjustment for unexpected conditions, and energy-efficient navigation. Deploying AI models directly on the robot hardware ensures low-latency, robust, and autonomous inspection even with limited communication.

## Customer Validation & Market Interest

Our intelligent inspection solution has been successfully validated through pilot projects with four Korean municipalities: Jindo, Siheung, Jeju, and Jeonju. Municipalities expressed high satisfaction with the robot's precise navigation, accurate defect detection, and efficient data acquisition.

These large-scale deployments yielded over 110,000 high-quality multimodal datasets (images, vibration, acoustic, IMU data), which were independently validated by TTA (Telecommunications Technology Association of Korea) for accuracy, consistency, and completeness, enabling advanced AI model development. Due to this proven performance, we are preparing overseas PoC projects in Canada, the United States, and the Middle East. Singapore's DSTA has requested

an on-site PoC, and a major Saudi Arabian company has requested a quotation for 200 kilometers of pipeline inspection. These domestic successes and international engagements demonstrate the technology's strong reliability, scalability, and global competitiveness.

## Robot Power & Endurance

Our inspection robot is essentially passive-powered, requiring minimal energy during standard operation. When moving forward, a canopy mechanism deploys to harness natural water flow, allowing travel without propulsion energy consumption. The propeller activates only for reversing or adjusting position, significantly reducing overall power consumption. Despite an internal battery, energy usage is extremely low, resulting in exceptionally long runtime without frequent recharging. This efficient design enhances operational endurance and reduces maintenance complexity.

## Technical Challenges & Mitigation

A key technical challenge in pipeline inspection is real-time wireless communication, often blocked by thick, reinforced steel pipes. While acknowledging current limitations, we are collaborating to advance communication technologies. However, Morphing-BOT's on-device AI enables autonomous decision-making and operation without external communication, allowing it to detect anomalies, adjust paths, and execute basic fault recovery. Expanding to oil and gas pipelines introduces further challenges, such as different fluid dynamics and higher pressures.

We are addressing these by partnering with Digital Petroleum, a major Saudi Arabian pipeline management company, to develop enhanced communication solutions and customized inspection protocols for complex environments.

**Partners**

Some of our partners in this project are listed below.

- AQUADATA (Canada): Collaboration for North American pipeline inspection and PoC deployments.
- Digital Enterprise (Qatar): Strategic partner for Middle East expansion, providing access to major water and infrastructure projects.
- A**** (Saudi Arabia): Engagement in large-scale pipeline inspection and maintenance projects.
- Local Governments (Korea - Jeju, Jin-do, Jeonju, Siheung): Pilot deployment sites to validate technology and gather inspection data.
- University Collaborations (Korea): R&D partnerships with top-tier universities (Seoul National University, etc) for AI model development and sensor fusion research.
- Standford University, ENIT – International R&D

## Table 1: Summary of International Agreements

| Type | Date | Country | Signed |
|---|---|---|---|
| LOI | '25.05.27 | US | Standford University Loi |
| LOI | '25.05.27 | TUNIS | Ecole Nationale d'Ingénieurs de Tunis LoI |
| NDA | '25.04.07 | US | Qualcomm NDA |
| LOI | '25.03.10 | Canada | Aqua Data LoI |
| LOI | '25.02.26 | Qatar | MBK Holdings Digital Enterprises LoI |
| MOU | '25.02.26 | Qatar | MBK Holdings Digital Enterprises MOU |
| MOU | '25.02.23 | Qatar | FaB LAB MOU |
| MOU | '25.01.08 | US | GOOD DEVELOPEMENT GROUP MOU |
| MOU | '25.01.08 | US | USAKO MOU |
| MOU | '24.10.09 | Vietnam | DTS MOU |
| MOU | '24.10.09 | Vietnam | EG TECH MOU |
| MOU | '24.10.09 | Vietnam | ONEBIM VIETNAM CO., LTD. MOU |
| MOU | '24.11.05 | Saudi | NAVIGATE SAUDI MOU |
| MOU | '24.11.03 | Saudi | Digital Petroleum MOU |
| MOU | '24.11.05 | Qatar | TDH MOU |
| MOU | '24.10.24 | Tunusia | Idaraty MOU |
| MOU | '24.10.24 | Tunisia | Namla MOU |

## 2.2 Benefits of the use case

- Early anomaly detection reduces catastrophic pipeline failures and repair costs.
- Real-time monitoring enhances operational efficiency and safety.
- Predictive maintenance prolongs infrastructure lifespan and optimizes resource allocation.
- Scalable solution for global adoption in urban cities facing infrastructure aging challenges.

## 2.3 Future Work

Plans include enhancing the Morphing-BOT platform with improved real-time monitoring, integrating predictive maintenance AI models, and expanding to oil and gas pipelines. Next steps involve deploying additional pilot projects in North America and the Middle East to validate system scalability.

### Required Resources

Achieving these goals necessitates advanced AI model optimization tools, upgraded sensor modules for broader environmental adaptability, and enhanced cloud infrastructure for large-scale data processing. We also plan to expand data collection to diversify AI training datasets, improving asset condition prediction accuracy.

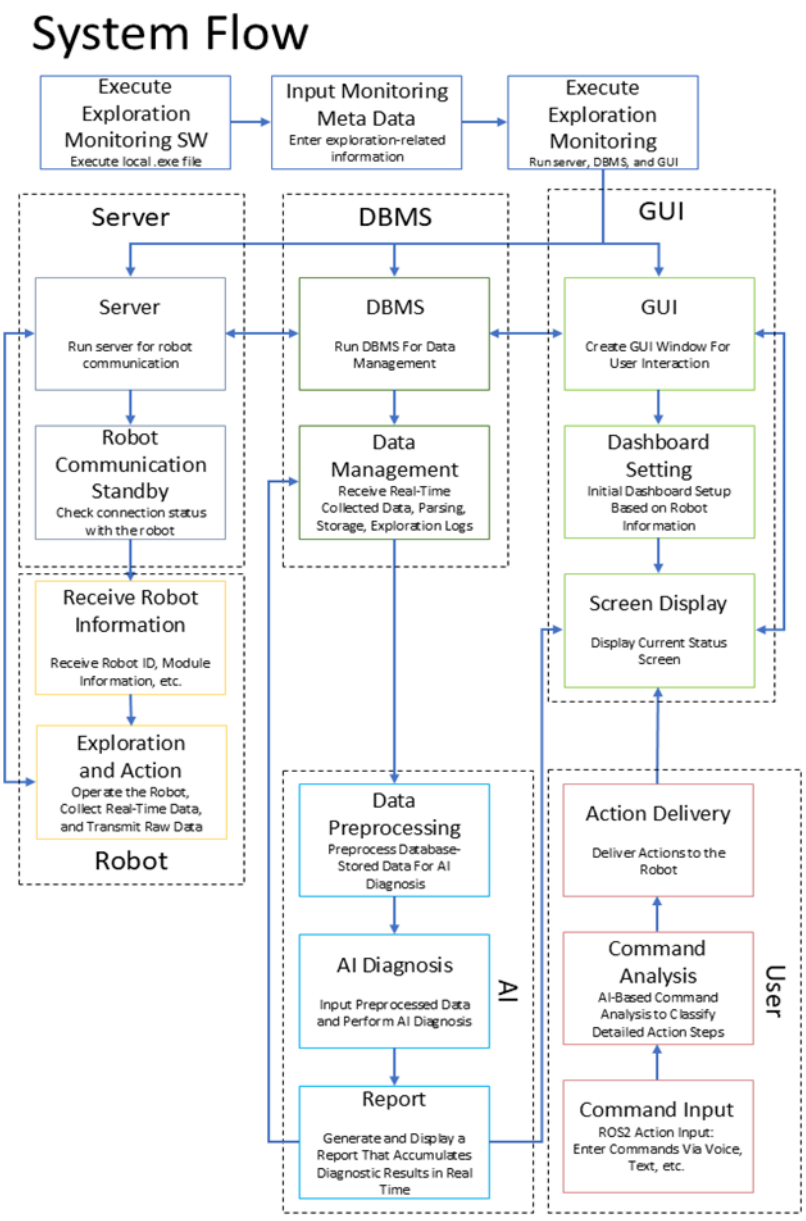
### Strategic Partnerships & Market Growth

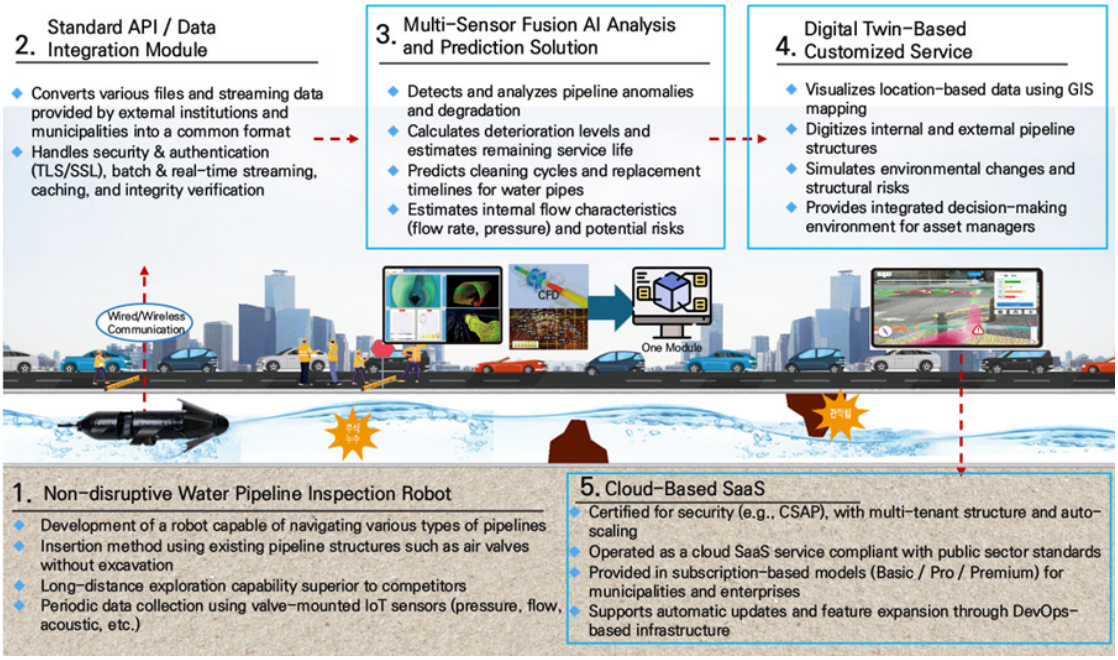
We are seeking strategic collaborations with global EPC companies, municipal water authorities, and smart city solution providers. Additionally, plans include establishing regional R&D hubs in the U.S. and the Middle East to support localized customization and service expansion.

# 3 Use Case Requirements

- **REQ-001**: It is critical that the system enables deployment of AI-powered robotic inspection systems capable of navigating narrow, curved, and buried pipelines.
- **REQ-002**: It is critical that the system supports integration with digital twin platforms to provide real-time data visualization and predictive maintenance capabilities.
- **REQ-003**: It is expected that the robotic device maintains an IP68 waterproof rating and modular hardware design to ensure adaptability across diverse and challenging pipeline environments.
- **REQ-004**: It is expected that the system operates on a secure and scalable cloud infrastructure for efficient data storage, processing, and analysis of multi-modal inspection data.
- **REQ-005**: It is of added value that the system complies with relevant international certifications (e.g., GCC certification, NSF standards) to facilitate global market deployment and regulatory acceptance.
- **REQ-006**: It is critical that the system provides real-time anomaly detection and predictive analytics through AI-based multi-sensor fusion, enabling early fault detection and operational decision-making.

# 4 Sequence Diagram






## 5 References

[1] TTA 대표홈페이지(영문)." Accessed: Jun. 09, 2025. [Online]. Available: https://www.tta.or.kr/eng/index.do#page2

[2] (Morphing I) Morphing BOT SOLUTION Introduction Eng - YouTube." Accessed: Jun. 09, 2025. [Online]. Available: https://www.youtube.com/watch?v=4yh_rq1R7Os

# Use Case 2: AI-based Semiconductor Design Automation and Optimization

**Organization**: AgileSoDA

**Country**: South Korea

**Contact Person(s)**:

Steve Kim, steve.kim@agilesoda.com

SeungYeol Baek, sybaek@agilesoda.ai

## 1 Use Case Summary Table

| Item | Details |
|---|---|
| Category | Manufacturing: Semiconductor Industry, AI Automation, EDA (Electronic Design Automation) |
| Problem Addressed | Rapidly increasing SoC (System-on-Chip) demand, shortage of design professionals, and labor-intensive design processes pose global challenges in sustainable industrial growth. Lengthy design cycles and manual verification hinder productivity and competitiveness. |
| Key Aspects of Solution | Reinforcement Learning-based AI for automated chip placement optimization<br>Seamless integration with existing tools and workflows<br>Multiple simulation capabilities<br>Expandable semiconductor design platform |
| Technology Keywords | AI, Reinforcement Learning, EDA, Semiconductor Design, Physical Layout, Macro Placement, Automation, ChipNSoDA[1] |
| Data Availability | Private (confidential semiconductor design data in LEF/DEF formats) |
| Metadata (Type of Data) | Semiconductor design files, physical layout data, macro cell placement data (LEF/DEF format) |
| Model Training and Fine-Tuning | Design of Graph Neural Network with appropriate nodes and edges, and train the model using prepared data with reinforcement learning.<br>Test performance using proximate cost metrics, verify accuracy with commercial EDA tools, and adjust parameters to improve performance. |
| Testbeds or Pilot Deployments | An internal pilot deployment was carried out in collaboration with our partner ASICLAND[2]. |

# 2 Use Case Description

## 2.1 Description

The semiconductor industry faces significant challenges due to rapidly expanding SoC demand, workforce shortages, and labor-intensive design processes[3]. Traditional physical design methods often require weeks or months of effort from specialized engineers, creating bottlenecks in production pipelines. As system semiconductor complexity increases, conventional approaches struggle to scale effectively, threatening competitiveness in a fast-moving market.

This use case aims to revolutionize semiconductor design by automating critical physical design steps that currently rely on extensive manual intervention. It leverages Reinforcement Learning (RL) for its effectiveness in navigating complex, high-dimensional design spaces and identifying optimal solutions for semiconductor physical layout optimization[4].

Specific objectives include:

- Shortening design time from weeks/months to hours.
- Reducing human error in complex layout decisions.
- Alleviating workforce constraints in specialized design roles.
- Enabling teams to handle more intricate layouts (hundreds of macro cells, thousands of standard cells).
- Fostering a more resilient design ecosystem, promoting innovation and broader access to advanced techniques.

The solution leverages RL-based artificial intelligence to optimize semiconductor physical design processes. Key technical components include:

- RL algorithms that learn optimal placement strategies for large-scale chip layouts
- Metadata Transformation Modules to protect sensitive design information
- Integration Layers ensuring seamless compatibility with existing design flows and tools

### Proven Performance

Initial deployments at partner companies yielded highly positive customer feedback, demonstrating significant design cycle reductions. An average optimization benefit of approximately 19% was observed. Specifically, for high-density designs, overall duration decreased by 20%, and the floor planning stage saw a remarkable 76% reduction, cutting design time from over a week to roughly one day.

### Broad Applicability & Future Direction

The AI-based floor planning solution is versatile, applicable across various designs by utilizing physical data from LEF/DEF files, rather than being limited by specific architectural types like RISC/CISC. Its AI focuses on universal optimization goals like minimizing wire lengths and congestion. Additionally, plans are underway to open-source the RL model and UI components, with the detailed scope depending on ongoing research and market demand.

This approach addresses immediate market pressures for production efficiency and long-term strategic needs, including talent shortages and technological advancement in the semiconductor industry. It is also expected to create high-quality jobs, enhance job satisfaction, and stimulate innovation across manufacturing sectors by improving semiconductor supply.

**Use Case Status:** Pilot Deployment

**Partners:** ASICLAND (Data Provider)

A partnership with ASICLAND brings access to reliable and high-quality data, ensuring that the RL-based solution can be trained effectively and deliver significant productivity gains.

## 2.2 Benefits of the use case

As design engineers are freed from low-value work, they can focus on higher-value, skill-building tasks, which elevates job quality and drives revenue growth. This shift creates more specialized employment opportunities and enhances professional satisfaction, ultimately developing a more robust and inclusive labor market. Moreover, by simplifying complex processes, AI enables even early-stage engineers and newcomers to enter the semiconductor design field—broadening access not only for individuals but also for emerging economies and developing nations. Increased semiconductor supply capacity—fueled by improved productivity—can further stimulate the development of various AI services, ultimately contributing to broader, sustainable economic growth.

By lowering entry barriers and attracting a new wave of talent, the adoption of AI is fueling growth in the semiconductor industry. This technological advancement streamlines the design process, significantly boosting production speed and strengthening supply capabilities. As accelerated semiconductor supply modernizes infrastructure and supports sustainable practices, AI delivers transformative innovation across a wide range of global manufacturing sectors.

## 2.3 Future Work

Data Collection Technology

- Expand design data collection through partnerships with additional design houses
- Collect and process data across the entire semiconductor design process
- Integrate verification metrics (e.g., design revision history, debug logs) for AI agent-based analysis

Model Development

- Develop AI agent-based project analysis models combining reinforcement learning with large language models
- Extend AI optimization to other semiconductor design stages (routing, timing optimization, power verification)

Create New Variations/Extensions

- Extend AI solution beyond macro placement to other design tool modules
- Develop project management and collaboration support features for design teams

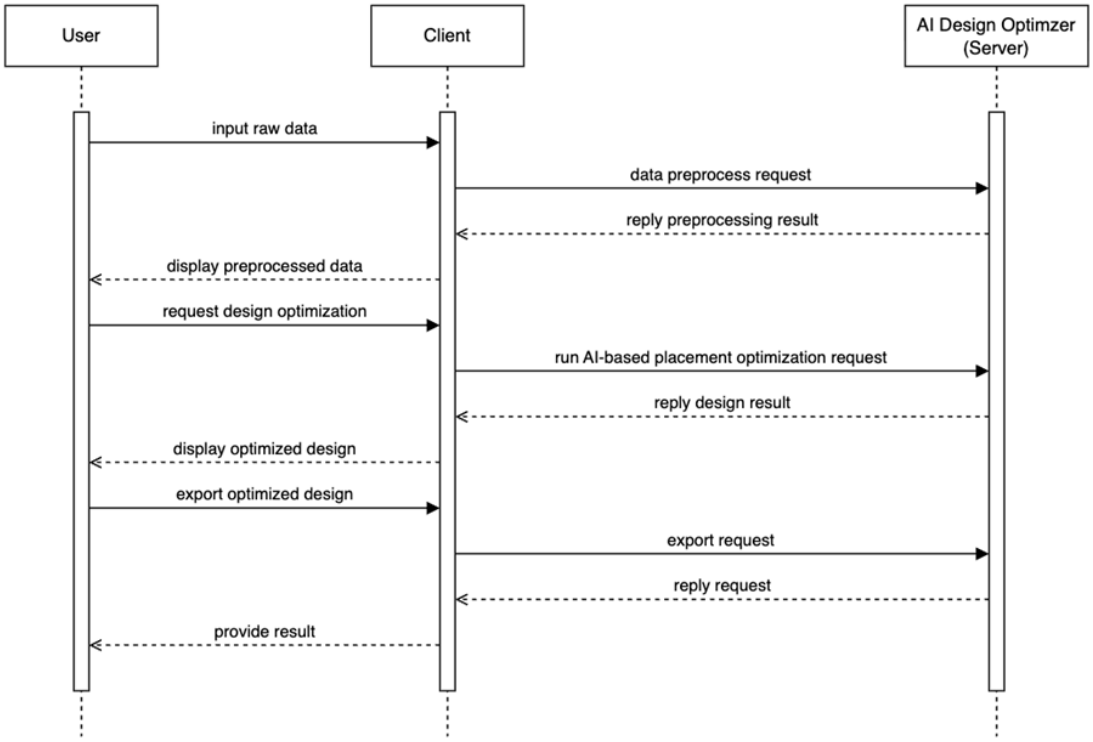
Semiconductor Design and Project Operation

- Implement AI agent monitoring and analysis throughout the semiconductor project lifecycle
- Develop early warning systems for project risks (workforce, schedule, process) to minimize redesign work

## 3 Use Case Requirements

- **REQ-01**: It is critical that the solution integrate seamlessly with existing design tools and workflows through a user-friendly interface, ensuring compatibility with established development environments and minimizing disruption to current practices.
- **REQ-02**: It is critical that the system support multiple simulation types to enable testing of various design scenarios based on specific project characteristics and requirements.
- **REQ-03**: It is of added value that the system lowers the technical entry barrier for novice designers while simultaneously maximizing the capabilities available to experienced professionals, thereby broadening the user base and enhancing overall design quality.
- **REQ-04**: It is expected that the system automates chip placement processes to reduce design time from the current weeks/months timeframe to hours, enabling rapid prototyping and development cycles.

## 4 Sequence Diagram



## 5 References

[1] "AgileSoda - ChipNSoDA." Accessed June 9, 2025. Available online: http://www.agilesoda.ai/sub/chip_soda.php

[2] "ASICLAND Co., Ltd." Accessed June 9, 2025. Available online: https://eng.asicland.com/

[3] S. Haque, Z. Eberhart, A. Bansal, and C. McMillan, "Semantic Similarity Metrics for Evaluating Source Code Summarization," in *IEEE International Conference on Program Comprehension*, IEEE Computer Society, 2022, pp. 36–47. https://doi.org/10.1145/nnnnnnn.nnnnnnn.

[4] A. Mirhoseini et al., "A Graph Placement Methodology for Fast Chip Design," *Nature*, vol. 594, no. 7862, pp. 207–212, June 2021. https://doi.org/10.1038/S41586-021-03544-W.

# Use Case 3: Intelligent Transmission Section Flow Control Agent in Power Systems

**Organization:** China Southern Power Grid Artificial Intelligence Technology Co., Ltd.

**Country**: China

**Contact Person(s)**:

Zhen Dai, zhendai07@foxmail.com or daizhen@csg.cn

## 1 Use Case Summary Table

| Item | Details |
|---|---|
| Category | Intelligent manufacturing |
| Problem Addressed | Current dispatcher's empirical adjustments introduce subjectivity, limit scenario complexity, prolong response times, and hinder multi-objective optimization. Growing renewable uncertainties exacerbate these challenges in accessing elevated or enclosed areas. |
| Key Aspects of Solution | This case applies deep reinforcement learning to enhance power system dispatching, focusing on reducing transmission section flows while maintaining safety constraints. |
| Technology | Reinforcement learning, Multi-Layer Perceptron (MLP)'s high-dimensional state modelling, Deep Deterministic Policy Gradient (DDPG)'s continuous control |
| Keywords | Transmission Section Flow, reinforcement learning |
| Data Availability | Data is private |
| Metadata (Type of Data) | Measurements in power system operations (e.g., switch states, active power, voltage magnitude, etc) and system parameters (e.g., generation and transmission limits). |
| Model Training and Fine-Tuning | The integration of MLP's high-dimensional state modelling with DDPG's continuous control enables adaptive section regulation through dynamic pattern learning while preserving stability constraints. |
| Testbeds or Pilot Deployments | Deployed in Southwest China's renewable-rich regional grid, the intelligent agent dynamically adjusts generation outputs in response to real-time fluctuations, consistently outperforming human dispatchers in comprehensive control evaluations. |
| Code repositories | Not Available |

# 2 Use Case Description

## 2.1 Description

### Intended Use

This case applies deep reinforcement learning to enhance power system dispatching, focusing on reducing transmission section flows while maintaining safety constraints.

### Problem to Solve

Current dispatcher's empirical adjustments introduce subjectivity, limit scenario complexity, prolong response times, and hinder multi-objective optimization. Growing renewable uncertainties exacerbate these challenges.

Limitations of Existing Solutions

Traditional methods are constrained by dispatchers' inconsistent expertise, multi-objective control complexity in multivariable optimization, and real-time operational time constraints, resulting in delayed and suboptimal decision-making.

### AI Methodology

The integration of MLP's high-dimensional state modeling with DDPG's continuous control enables adaptive section regulation through dynamic pattern learning while preserving stability constraints.

Benefits of the AI-Based Approach

- Efficiency: Decision-making processes can be accelerated from minute level decisions to second.
- Versatility: More complex scenarios can be generalized.
- Multi-Objective: System balance, transmission limits, renewable energy integration, and operational costs can be simultaneously optimized through coordinated controls.
- Scalability: Evolving grid complexities and constraints can be dynamically accommodated through AI-driven adaptation mechanisms.

### Drawbacks of the AI-Based Approach

- Safety Compliance: AI systems may compromise safety requirements under generation/load uncertainties.
- Model Robustness: Incomplete knowledge of physics may reduce model robustness and generalization capabilities.
- Data Dependency: Inaccessibility of high-quality data for model training demands may limit practical application effect.

By embedding physical knowledge into AI models, these limitations can be addressed, enhancing the robustness and reliability of the system.

**Use Case Status:** The use case is part of a larger product development

## 2.2 Benefits of the use case

1. Improving Energy Efficiency. Through real-time optimization of generating unit outputs, the Intelligent Section Flow Agent ensures that coal-fired power plants operate at their most efficient levels. This reduces coal consumption, minimizes carbon emissions, and promotes cleaner energy usage compared to manual scheduling methods.
2. Enhancing Economic Efficiency of Power Grid Operations. The Intelligent Section Flow Agent optimizes control strategies to maintain transmission flows within safe and efficient limits. By reducing the need for costly investments in transmission infrastructure and lowering operational expenses, the agent ensures cost-effective power grid operations, which enhances economic productivity and stability.
3. Facilitating Renewable Energy Development. The agent ensures the stable operation of long-distance transmission networks, facilitating the efficient delivery of renewable energy from resource-rich regions in southwestern and northwestern China to load centers in the southeast. By supporting renewable energy integration, the Intelligent Section Flow Agent accelerates the transition to a sustainable energy system and reduces reliance on fossil fuels.

## 2.3 Future Work

Our future work will focus on refining and extending the Intelligent Section Flow Agent system to address emerging challenges in power grid operation. The enhancements will include architectural improvements, advanced modeling techniques, and practical deployment strategies to maximize system robustness, adaptability, and scalability.

### 1. Enhanced Data Management and Integration

Future developments will emphasize integrating additional data sources, such as high-temporal-resolution measurements and environment information, to enrich real-time monitoring. Improved preprocessing algorithms will be developed to handle diverse data formats, enhance quality, and ensure greater accuracy for downstream decision-making processes.

### 2. Advanced Agent Training

We plan to explore hybrid learning approaches by incorporating physical knowledge into reinforcement learning models. This integration will enhance model generalization, enabling the Intelligent Section Flow Agent to handle more complex and extreme scenarios while maintaining grid safety and efficiency. Moreover, transfer learning techniques will be utilized to accelerate agent training across similar operational regions, reducing computational costs.

### 3. Distributed and Scalable Architecture

The scheduling platform will be optimized for greater scalability through enhanced distributed computing frameworks. This will ensure reliable deployment in large-scale, multi-regional power grid systems.

### 4. Real-Time Optimization and Emergency Handling

Future iterations will incorporate adaptive algorithms to improve the agents' real-time response capability to dynamic grid changes and emergencies. Advanced risk assessment modules will

be integrated to predict and mitigate potential failures, further enhancing grid stability and resilience.

**5. Interoperability and Industry Adoption**

We will work on designing standardized APIs and protocols to ensure seamless compatibility with existing power grid dispatch systems. Pilot programs will be initiated in collaboration with regional power utilities to validate system performance under real-world conditions. Feedback from these implementations will guide iterative system improvements.

By focusing on these areas, the Intelligent Section Flow Agent system will evolve into a robust, intelligent, and scalable solution, driving innovation in power grid management and supporting the sustainable development of energy systems.

## 3 Use Case Requirements

Hardware Requirements:

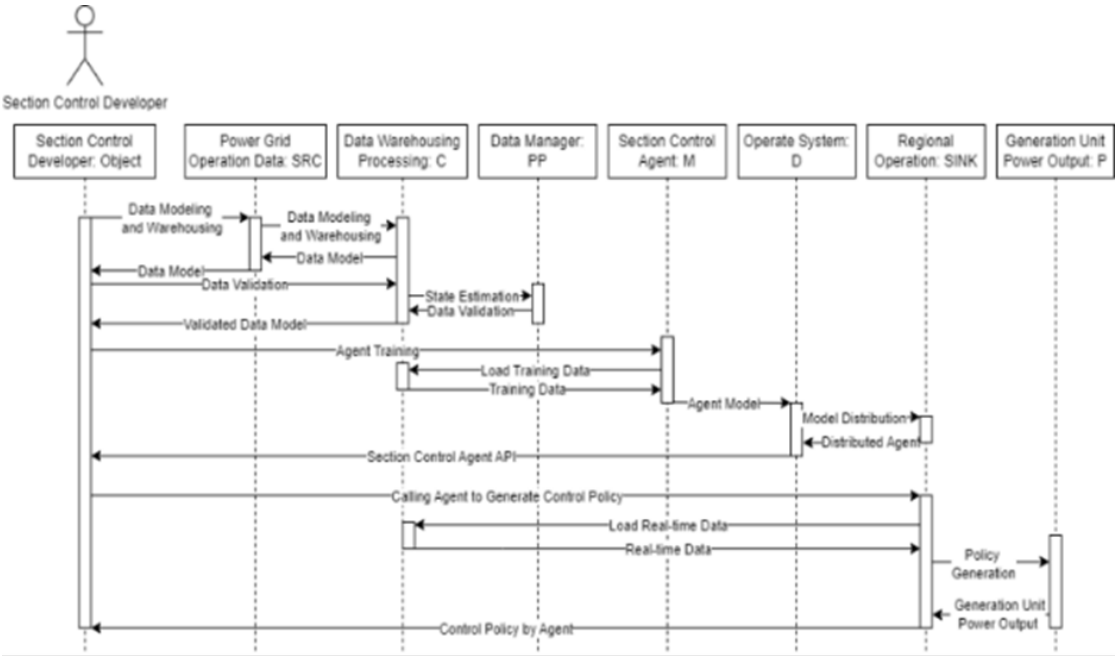
For development/training: Minimum 16-core CPU + 64GB RAM (Recommended: 32-core CPU + 128GB RAM + GPU acceleration). One core acts as the central node, while the remaining cores serve as edge computing nodes. Distributed training requires multiple Linux nodes with ≥10GbE high-speed network. Deployment needs edge devices supporting Python inference (Docker containerization recommended), with industrial PCs for real-time control scenarios.

Environment Requirements:
Python ≥3.7 (3.8/3.9 recommended) with virtual environment dependency management. Core software stack:

- ray[rllib] (distributed RL framework)
- torch/tensorflow (deep learning backend)
- gymnasium (environment interface)
- pandas/numpy (data processing)

Power system simulation requires power flow simulators, with tensorboard recommended for real-time monitoring.

## 4 Sequence Diagram




## 5 References

[1] Li, Peng, et al. "A novel graph reinforcement learning approach for stochastic dynamic economic dispatch under high penetration of renewable energy." 2022 4th Asia Energy and Electrical Engineering Symposium (AEEES). IEEE, 2022.

[2] Yan J, Li Y, Yao J, et al. Look-Ahead Unit Commitment With Adaptive Horizon Based on Deep Reinforcement Learning[J]. IEEE Transactions on Power Systems, 2024, 39(2): 3673-3684.

# Use Case 4: AI-Driven Virtual Try-On and Garment Generation System for Sustainable Indian Fashion

**Organization**: Indian Institute of Information Technology Sonepat

**Country**: India

**Contact person**:

Dr. Vinay Kaushik, vkaushik@iiitsonepat.ac.in, +91 9013241634

## 1 Use Case Summary Table

| Item | Details |
|---|---|
| Category | Manufacturing |
| Problem Addressed | Both designers and online shoppers struggle to visualize how outfits, ranging from sarees and sherwanis to contemporary dresses and suits, will appear on real bodies, resulting in high return rates. Additionally, a lack of demand forecasting leads to overproduction and inventory waste. This creates significant environmental and financial costs for brands and artisans alike. |
| Key Aspects of Solution | 1. Using cutting-edge generative AI models like Stable Diffusion and Transformer-based networks, users can digitally try on garments with realistic simulation of fabric drape, fit, and style.<br>2. AI-powered demand forecasting models will analyze trends and consumer behavior to inform what should be designed or produced, minimizing waste and aligning supply with actual demand. |
| Technology Keywords | Generative AI, diffusion models, virtual try-on, demand forecasting system, garment generation |
| Data Availability | Public dataset (Apparel Classification for Indian Ethnic Clothes) found at [1] |
| Metadata (Type of Data) | Photos, text prompts, sketches |

(continued)

| Item | Details |
|---|---|
| Model Training and Fine-Tuning | Pose-Agnostic Garment Transfer: Uses Stable Diffusion to warp garments realistically across body types.<br>Contrastive Language-Image Pre-training (CLIP)-Guided Fine-Tuning: Preserves traditional textile patterns (Banarasi, Kanjeevaram, etc.) during virtual try-ons.<br>Trust Mechanism: AI models incorporate real-world garment physics (wrinkles, elasticity) to build user confidence in how a garment will fit and feel. Fit scoring and visual comparison tools increase alignment with actual products.<br>StyleGAN3 + ViT (Vision Transformer): Creates high-fidelity textures for Indian fabrics while retaining weave details.<br>Diffusion Priors for Fabric Simulation: Models wrinkles, elasticity, and physics-aware garment movement.<br>Text-to-Design (Imagen, Parti): Enables on-demand, AI-generated garments, blending modern cuts with traditional motifs.<br>Transformer-based Forecasting Models: Predict consumer preferences based on trends, demographics, regional festivals, and past behavior. |
| Testbeds or Pilot Deployments | Looklet dressing room demo [2]<br>Virtual makeup demo [3]<br>Outfit Anyone: Ultra-high quality virtual try-on for Any Clothing and Any Person [4]<br>Outfit Anyone in the Wild: Get rid of Annoying Restrictions for Virtual Try-on Task [5] |
| Code repositories | IDM-VTON: Improving Diffusion Models for Authentic Virtual Try-on in the Wild dataset [6]<br>CatVTON: Concatenation Virtual Tryon dataset [7]<br>Multimodal-Conditioned Latent Diffusion Models for Fashion Image Editing (Garment Design) [8]<br>Generative Pretrained Hierarchical Transformer for Time Series Forecasting (Forecasting Modeling) [9] |

# 2 Use Case Description

## 2.1 Description

The Indian fashion industry blends traditional craftsmanship with modern designs, serving diverse consumer preferences across regions. However, both designers and online shoppers struggle to visualize how outfits, ranging from sarees and sherwanis to contemporary dresses and suits, will appear on real bodies, resulting in high return rates. Additionally, a lack of demand forecasting leads to overproduction and inventory waste. This creates significant environmental and financial costs for brands and artisans alike.

To address these challenges, we propose an AI-driven virtual try-on, garment generation, and demand forecasting system. Using cutting-edge generative AI models like Stable Diffusion and Transformer-based networks, users can digitally try on garments with realistic simulation of fabric drape, fit, and style. AI-generated garment designs based on user inputs, such as photos, text prompts, or sketches, support both consumer personalization and designer creativity.

Further, AI-powered demand forecasting models will analyze trends and consumer behavior to inform what should be designed or produced, minimizing waste and aligning supply with actual demand. This also empowers traditional weavers by digitizing and adapting their patterns for wider markets.

Users will interact with the system through multiple interfaces: Mobile apps (upload photos or avatars), In-store kiosks (try-ons in retail or artisan fairs), and Designer consoles (for sketch or text-based garment design). The result is a highly personalized, trust-driven, and sustainable ecosystem for Indian fashion.

Problem it Aims to Solve

High Return Rates: Inaccurate sizing and fit perception in online shopping contribute to returns, increasing waste and logistics costs.

Overproduction & Waste: Uncertainty in consumer demand leads to excessive production, harming sustainability.

Trust Gap in Virtual Try-On: Users may not trust the virtual version of garments to match their physical fit and feel.

Limited Access to Traditional Fashion: Local artisans and weavers struggle to reach a digital audience, limiting their market opportunities.

Consumer Decision Fatigue: Shoppers face difficulty selecting garments that suit their preferences and body types.

Limitations of Existing Solutions

Existing try-ons lack realistic draping, texture mapping, and adaptability to body types.

Consumers rely on inconsistent sizing charts, leading to mismatches.

Current e-commerce platforms offer static catalogs with minimal customization options. Designers do not have an AI tool which can take real world design attributes as input and generate a garment image virtually. A garment generation and virtual tryon solution tailored to India's use case does not exist.

Benefits and Drawbacks of AI-Based Approach

Benefits:

- AI-driven diffusion models provide high-quality, dynamic visualizations of clothing.
- AI suggests designs tailored to user preferences.
- Reduced returns and optimized production contribute to environmental conservation.
- Digitization of traditional designs expands market access.
- Helps designers look how their imagined design will look in the real world even before any production begins.

Drawbacks:

1. Requires high-end processing power for real-time rendering.
2. Storing user images raises ethical and security challenges.

3. Small-scale retailers may face difficulties integrating AI solutions.

This AI-powered system revolutionizes Indian fashion retail by bridging traditional and modern fashion, enhancing user experience, and promoting sustainability.

## 2.2 Benefits of the use case

1. Addresses key sustainability challenges in India's modern and traditional fashion industry. India's fashion sector, with its mix of handcrafted textiles and mass-produced apparel, struggles with high return rates, overproduction, and excess textile waste. Virtual try-on technology powered by generative AI models (such as Stable Diffusion for hyper-realistic visualizations and Transformer-based AI for personalized garment design) can significantly reduce waste by helping consumers make more informed purchase decisions. AI-driven demand forecasting helps brands better anticipate consumer preferences and produce accordingly, reducing overstock and textile waste.
2. AI-powered garment generation helps revive traditional weaving techniques by enabling digital customization of handcrafted designs. By digitizing and preserving cultural craftsmanship, the proposal enhances economic opportunities for artisans and helps designers make the right garment design decision,
3. Climate action benefits from AI's ability to reduce return-related logistics emissions and minimize excess textile production, both of which contribute to carbon emissions. AIs predictive capabilities allow brands to manufacture only what is needed, cutting down industrial waste and energy usage.

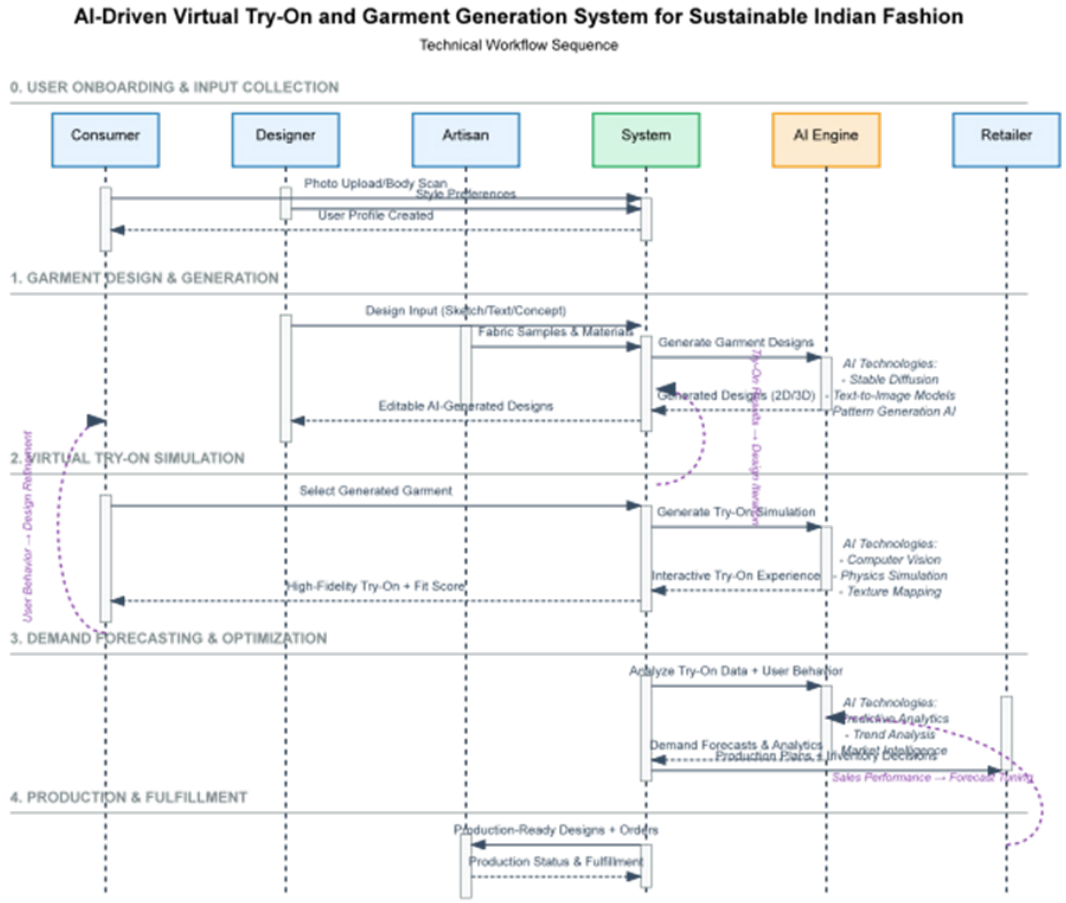
## 2.3 Future Work

1. Data collection: Curating a high-quality, diverse dataset is foundational for robust model performance. This includes collecting images of different types and styles of garments, traditional textile patterns (e.g., Banarasi, Kanjeevaram, Chikankari, Baluchari, Chanderi), user poses across body types, backgrounds, and consumer preference data. Attention must be paid to ensuring demographic and cultural representation, and to ethical data practices including consent and anonymization.
2. Proof of concept development: A functional PoC will demonstrate key capabilities such as virtual try-on accuracy, garment texture preservation, and generative design from text or sketches. This phase will validate the technical feasibility, user experience, and business impact, using limited-scale deployments via mobile apps or kiosks at artisan fairs or retail pilots.
3. Model development: Core components include: (1) generative models for garment synthesis, (2) embedding models for texture consistency, (3) physics-aware simulation modules for realistic draping, and (4) forecasting models for demand prediction. This stage also involves fine-tuning models on region-specific data and optimizing performance for real-time inference.
4. Create new variations/extensions to the same use case: The use case can evolve to support applications like personalized garment recommendations, person avatar integration for deeper realism, and virtual tailoring. Extensions could also focus on fashion education, enabling students and artisans to design and preview garments digitally before physical prototyping.
5. Standards development related to the use case: To ensure interoperability, fairness, and scalability, the initiative must contribute to the development of technical standards related to virtual try-on formats, AI-generated fashion metadata, and privacy-preserving mechanisms in user-data-driven AI. Collaborating with industry and regulatory bodies will be key to establishing the best practices for ethical and sustainable AI in fashion.

## 3 Use Case Requirements

- **REQ-001**: It is critical that AI-driven virtual try-on tools enable realistic trustworthy visualization of garments across diverse body types and poses, supporting seamless integration of traditional and modern design elements.
- **REQ-002**: It is critical that generative AI systems significantly accelerate the design process by enabling image generation from sketches, text prompts, or references, while preserving cultural and textile authenticity.
- **REQ-003**: It is critical that sustainability objectives are embedded in the system through AI-powered demand forecasting and waste reduction mechanisms that align production with consumer trends and regional preferences.

## 4 Sequence Diagram




## 5 References

[1] "IndoFashion: Apparel Classification for Indian Ethnic Clothes," [Online]. Available: https://indofashion.github.io/ (accessed: Jun. 15, 2025).

[2] Looklet, "demo / looklet dressing room," [Online]. Available: https://dressing-room.looklet.com/ (accessed: Jun. 15, 2025).

[3] Banuba, "AR Virtual Makeup Online Technology Demo," [Online]. Available: https://www.banuba.com/hubfs/virtual-makeup-demo/index.html (accessed: Jun. 15, 2025).

[4] Hugging Face, "Outfit Anyone: Ultra-high quality virtual try-on for Any Clothing and Any Person," [Online]. Available: https://huggingface.co/spaces/HumanAIGC/OutfitAnyone (accessed: Jun. 15, 2025).

[5] Hugging Face, "Outfit Anyone in the Wild: Get rid of Annoying Restrictions for Virtual Try-on Task," [Online]. Available: https://huggingface.co/spaces/selfit-camera/OutfitAnyone-in-the-Wild (accessed: Jun. 15, 2025).

[6] Github, "IDM-VTON," [Online]. Available: https://github.com/yisol/IDM-VTON (accessed: Jun. 15, 2025).

[7] Github, "CatVTON:Concatenation Is All You Need for Virtual Try-On with Diffusion Models ," [Online]. Available: https://github.com/Zheng-Chong/CatVTON (accessed: Jun. 15, 2025).

[8] Github, "Ti-MGD: Multimodal-Conditioned Latent Diffusion Models for Fashion Image Editing," [Online]. Available: https://github.com/aimagelab/Ti-MGD (accessed: Jun. 15, 2025).

[9] Github, "GPHT: Generative Pretrained Hierarchical Transformer for Time Series Forecasting," [Online]. Available: https://github.com/icantnamemyself/GPHT (accessed: Jun. 15, 2025).

[10] Pathak, Sanhita, Vinay Kaushik, and Brejesh Lall. "Single stage warped cloth learning and semantic-contextual attention feature fusion for virtual tryon." In 2024 IEEE International Conference on Multimedia and Expo (ICME), pp. 1-6. IEEE, 2024.

[11] Pathak, Sanhita, Vinay Kaushik, and Brejesh Lall. "MAC-VTON: Multi-modal Attention Conditioning for Virtual Try-on with Diffusion-Based Inpainting." In 2024 39th International Conference on Image and Vision Computing New Zealand (IVCNZ), pp. 1-6. IEEE, 2024.

[12] Zheng Chong, Xiao Dong, Haoxiang Li, shiyue Zhang, Wenqing Zhang, Hanqing Zhao, xujie zhang, Dongmei Jiang, Xiaodan Liang. "CatVTON: Concatenation Is All You Need for Virtual Try-On with Diffusion Models." In 2025 International Conference on Learning Representations (ICLR), OpenReview 2025.

[13] Choi, Yisol, Sangkyung Kwak, Kyungmin Lee, Hyungwon Choi, and Jinwoo Shin. "Improving diffusion models for authentic virtual try-on in the wild." In European Conference on Computer Vision, pp. 206-235. Cham: Springer Nature Switzerland, 2024.

[14] Baldrati, Alberto, Davide Morelli, Giuseppe Cartella, Marcella Cornia, Marco Bertini, and Rita Cucchiara. "Multimodal garment designer: Human-centric latent diffusion models for fashion image editing." In Proceedings of the IEEE/CVF international conference on computer vision, pp. 23393-23402. 2023.

[15] Zhiding Liu, Jiqian Yang, Mingyue Cheng, Yucong Luo, and Zhi Li. 2024. Generative Pretrained Hierarchical Transformer for Time Series Forecasting. In Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD '24). Association for Computing Machinery, New York, NY, USA, 2003–2013.

# Use Case 5: Explosion-Proof Intelligent Inspection Robot Solution

**Organization**: Sevnce Robotics Co., Ltd.

**Country**: China

**Contact person**:

Chengcheng Wang; wangchengcheng@sevnce.com

## 1 Use Case Summary Table

| Item | Details |
|---|---|
| Category | Manufacturing |
| Problem Addressed | In scenarios such as traditional manual inspection of power facilities and industrial equipment, there are issues like low efficiency, high safety risks, and incomplete data collection. Existing inspection robots have insufficient intelligence and struggle to adapt to complex environments and handle abnormal situations autonomously. |
| Key Aspects of Solution | Adopt embodied intelligence technology and combine multimodal sensors (such as infrared thermal imagers, high - definition cameras, gas detectors, etc.) to achieve environmental and equipment status perception. Utilize deep learning and reinforcement learning algorithms to endow robots with autonomous path planning, abnormal identification, and emergency response capabilities. Through edge computing and cloud - based data analysis, realize real - time monitoring of equipment status and predictive maintenance. |
| Technology Keywords | Embodied intelligence, multimodal sensors, deep learning, reinforcement learning, edge computing, predictive maintenance |
| Data Availability | Private. Data is sourced from sensor data collection in actual inspection scenarios, historical equipment operation data, etc. |
| Metadata (Type of Data) | Visual data (equipment appearance images), thermal imaging data (equipment temperature distribution), gas concentration data, vibration data, sound data |
| Model Training and Fine-Tuning | Based on a large number of equipment images and operation data, use deep learning models to train for abnormal feature identification. Optimize the robot's inspection path planning and decision - making model in complex environments through reinforcement learning. |
| Testbeds or Pilot Deployments | Pilot projects were carried out in a large - scale substation and a chemical plant. Link: |

# 2 Use Case Description

## 2.1 Description

In industrial fields such as power and chemical industries, equipment inspection is a crucial part of ensuring production safety and stable operation. Traditional manual inspection relies on the experience and subjective judgment of inspectors, which is not only inefficient but also poses great safety hazards in high - risk environments (such as high - voltage power facilities and toxic gas environments). Moreover, manual inspection is difficult to achieve high - frequency and full - coverage, and it is easy to overlook potential equipment failures.

This use case aims to develop an intelligent inspection robot based on embodied intelligence to realize the intelligent, automated, and efficient inspection of industrial equipment. The robot can autonomously complete tasks such as equipment appearance inspection, temperature monitoring, gas leak detection, and vibration analysis in complex and high - risk industrial environments, collect multi - dimensional data in real - time, and through intelligent analysis, promptly detect equipment abnormalities, provide early warnings of failures, and reduce equipment failure rates and downtime.

The project integrates multimodal sensors to build the robot's perception system, enabling it to comprehensively obtain equipment status information. Deep learning algorithms are used to analyze the collected data to accurately identify equipment abnormalities. Combined with reinforcement learning, the robot can autonomously plan the optimal inspection path according to environmental changes and respond to emergencies when encountering unexpected situations.

It is expected that this solution can increase the equipment inspection efficiency by more than 50%, reduce manual inspection costs by 60%, and decrease the probability of sudden equipment failures by 40%, effectively guaranteeing the safety and stability of industrial production and improving the enterprise's operation and management level.

Use Case Status: Pilot stage

### Partners

Industrial equipment manufacturing enterprises: Provide equipment operation data and technical support, and assist in the adaptation between the robot and the equipment.

University robot laboratories: Provide technical guidance in the research and development of embodied intelligence algorithms and model optimization, and jointly carry out scientific research cooperation.

## 2.2 Benefits of the use case

Promote the innovation of industrial inspection technology and facilitate the application of intelligent infrastructure in the industrial field.

Through predictive maintenance, reduce resource waste caused by equipment failures, improve production efficiency, and promote sustainable production.

Automated inspection technology reduces high - risk manual operations, ensures the safety of employees, and frees up human resources for more valuable work, driving industrial economic growth.

## 2.3 Future Work

Further optimize the robot's multimodal data fusion algorithm, develop a more intelligent embodied brain, and improve the accuracy and robustness of abnormal identification.

Expand the application scenarios of the robot and promote it to more fields such as oil and gas pipelines and railway facilities.

Strengthen cooperation with more industrial enterprises, collect data from different scenarios, and enrich the robot's training samples.

Required resources include high - performance edge computing devices, more types of high - precision sensors, and data interface development resources for cooperation with enterprises.

1) what kind of plug and play is supported?

e.g. can the user integrate different AI model trained based on DeepSeek or other models?

For example, can users integrate different Model A trained based on Deepseek or other models? Yes, different Model A can be integrated. We also have our own developed large - scale chemical safety model.

2) what kind of connectivity is needed? e.g. 5G?

Both 4G and 5G are available, and a local area network is also acceptable.

3) what are the types of sensors which are used? what kind of support is required for maintenance?

The sensors are made in China. Regarding maintenance, we can deliver the robot technology to on - site workers.

4) is it possible to describe any APIs supported? e.g. for application developers based on the solution

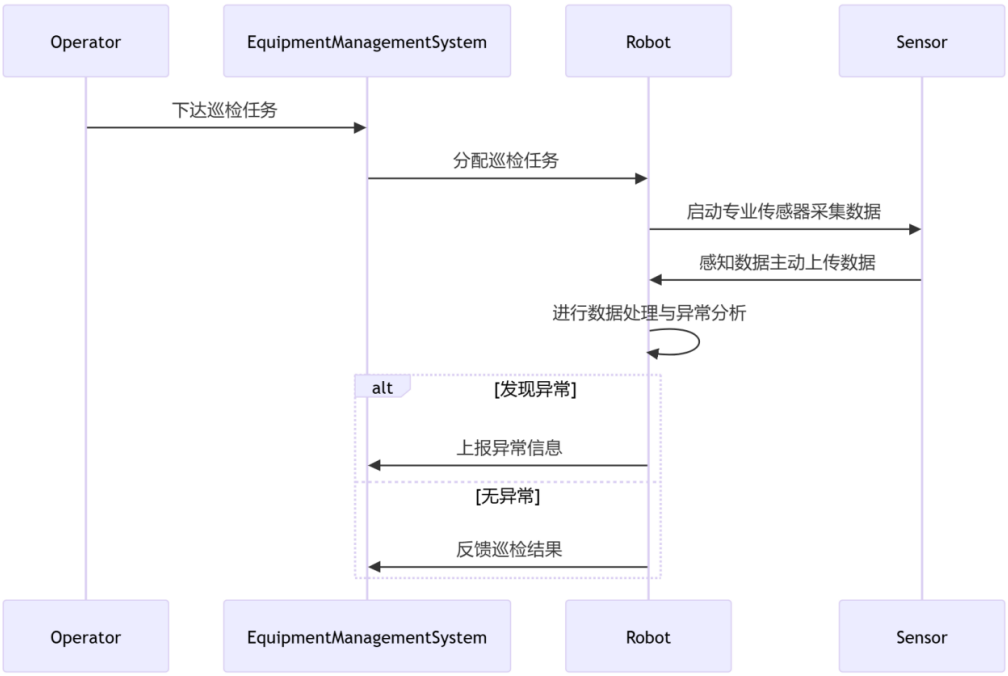
Robot open interfaces:

Login interface, user project permission interface, product pagination list (including device online status), robot failure information pagination, inspection task pagination list, inspection task details, inspection point details, inspection task result pagination, task result details, area pagination list, area details, main device pagination list, main device details, ledger pagination list, ledger details, warning pagination list, warning details, one - click warning processing, batch warning processing, query the warning list of the current task, ledger inspection result data, single inspection task result statistics, robot inspection data statistics, warning statistics, and statistics of the latest status of the inspection ledger.

## 3 Use Case Requirements

- **REQ - 01**: The robot should have multimodal perception capabilities, be able to work normally in complex lighting, dust, high - temperature, and other environments, and the data collection accuracy should meet the requirements of equipment detection.
- **REQ - 02**: The robot's abnormal identification response time should not exceed 5 seconds to ensure timely detection of equipment failures.
- **REQ - 03**: The robot's single - charge endurance time should be no less than 6 hours to meet the needs of long - term inspection tasks.
- **REQ - 04**: The robot needs to be seamlessly integrated with the enterprise's existing equipment management system to achieve data sharing and unified management.

## 4 Sequence Diagram




## 5 References

# Use Case 6: Generative AI-Driven Full-Process Computer-Aided Design

**Organization**: AInnovation Technology Group Co. Ltd.

**Country**: China

**Contact Person(s)**:

Cheng Yan, chengyan@ainnovation.com

Wang Fei, wangfei@ainnovation.com

## 1 Use Case Summary Table

| Item | Details |
|---|---|
| Category | Manufacturing |
| Problem Addressed | Traditional CAD software features complex interfaces and high usage thresholds, demanding stringent experience from operators. |
| Key Aspects of Solution | Industrial Multimodal LLMs powered, ChatCAD, Text-to-CAD, Image-to-CAD, Natural Language Driven, Design Requirements Understanding, Interactive Design, Proprietary Domain-Specific Language (DSL) Framework |
| Technology Keywords | Industrial Multimodal LLMs, DSL for Generative CAD, Synthetic Data, Supervised Fine-tuning (SFT), Retrieval-Augmented Generation (RAG) |
| Data Availability | Private |
| Metadata (Type of Data) | Texts and images |
| Model Training and Fine-Tuning | AInno-75B Industrial Multimodal LLM[1] fine-tuned with large-volume synthetic data in the CAD area, incorporating our proprietary DSL for precise representation of CAD elements. |
| Testbeds or Pilot Deployments | Integrated by the CAD software of Bentley Systems (a world-class CAD software provider) with MVP product development completed in Q4 2024. Currently being promoted to commercial clients for feedback collection to enhance generalization capabilities.<br>With 95% of F1 score |

## 2 Use Case Description

### 2.1 Description

Traditional CAD software features complex interfaces and high usage thresholds, demanding stringent experience from operators. AInnovation has introduced industrial large model

technology into the field of industrial design, launching applications like Text-to-CAD and Image-to-CAD under the name ChatCAD[2]. Our solution is built upon a custom-designed Domain-Specific Language (DSL) that effectively represents diverse CAD elements across both 2D and 3D environments[3]. This specialized DSL enables precise interpretation of geometric constraints, material properties, and functional requirements in generative CAD scenarios.

Through a simple conversational Q&A format, it swiftly understands designers' creative intentions, automatically generating industrial design drawings that meet requirements, and supports exporting to traditional industrial design software for manual fine-tuning. The DSL framework demonstrates strong adaptability across multiple CAD domains including PID process design, architectural planning, and mechanical component design, ensuring consistent performance in cross-domain applications.

This not only significantly enhances the efficiency and quality of industrial design while lowering the entry barrier but also promotes the extensive application of large model technology in upstream industrial production. Currently, this product can be applied in mechanical component design, reading engineering construction blueprint silhouettes, industrial product appearance design, electrical design, architectural drawing design, etc., markedly reducing the entry threshold for industrial design.

Use Case Status: The use case is part of a larger product development and successfully completed its MVP at the end of 2024, with plans to deploy business applications throughout 2025.

**Partners**

AInnovation partnered with Bentley Systems to explore the application of ChatCAD in the field of engineering design software, and help promote the intelligent transformation of China's infrastructure sector.

## 2.2 Benefits of the use case

ChatCAD helps enhance the efficiency and quality of industrial design, lowers the barriers to entry, drives the application of LLM technology in the upstream of industrial production, and accelerates the intelligent upgrading of industries. By leveraging AI technology, productivity can be enhanced, enabling designers to iterate design concepts faster, shorten product development cycles, and bring new products to market earlier. Additionally, as ChatCAD's applications continue to expand and deepen, it will spawn a series of new intelligent-technology-based products, services, and business models, opening up new growth opportunities for economic development

## 2.3 Future Work

The future plan aims to integrate generative AI technology throughout the entire process of computer-aided design (CAD) in one or more fields. This involves developing new industrial design software with three key functionalities: automatic conversion from images to CAD drawings (Image-to-CAD), autonomous review of CAD drawings (CAD-to-Text), and automatic generation of CAD drawings from text-based design requirements (Text-to-CAD).

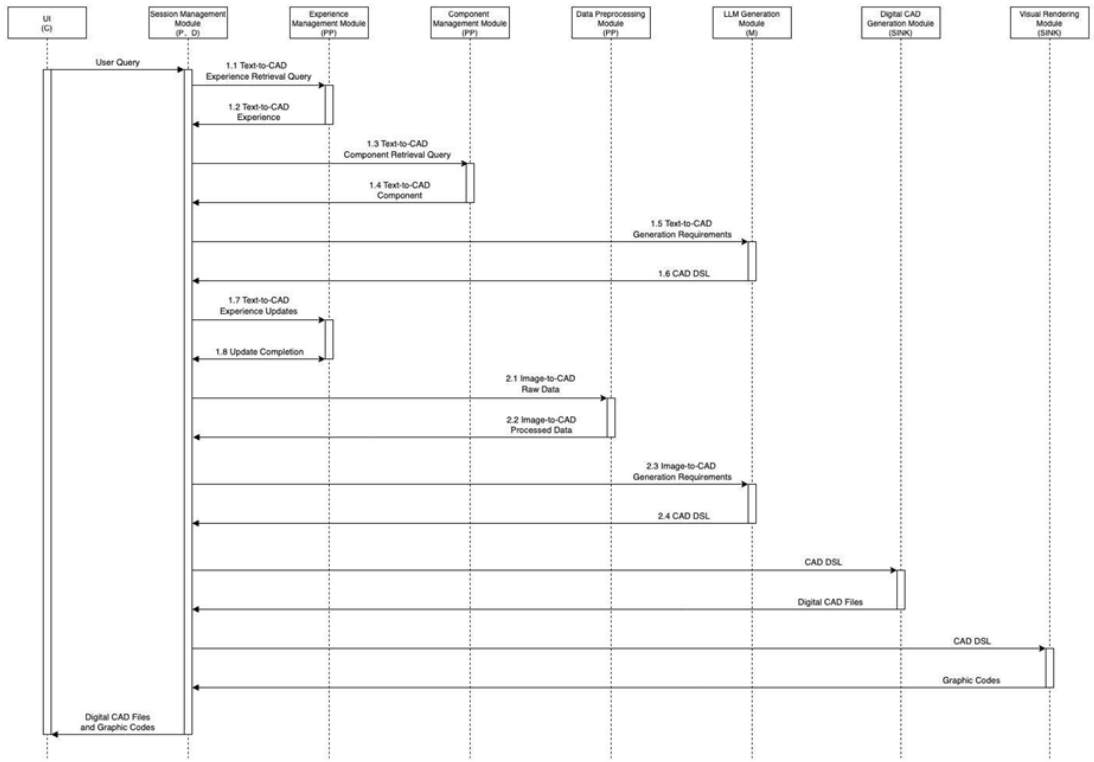
As individual functionalities mature, we plan to converge these capabilities into an integrated product ecosystem, enabling bidirectional translation between natural language, visual representations, and CAD models.

The goal is to achieve AI-driven 2D/3D CAD design capabilities including converting, reading, reviewing, and generating drawings. This initiative seeks to reduce the complexity of design operations, enhance design efficiency, and improve design quality.

## 3 Use Case Requirements

- **REQ-01**: It is critical that the system provides a user-friendly interface that allows users to interact with the system with natural language;
- **REQ-02**: It is critical that the system is capable of accurately understanding users' requirements (texts and images) for design;
- **REQ-03**: It is critical that the system is capable of generating digital CADs based on users' textual requirements;
- **REQ-04**: It is critical that the system is capable of extracting contents from input design images w/o textual instructions and generating digital CADs;
- **REQ-05**: It is critical that the system is capable of outputting editable CAD models in the common formats in the CAD area.

## 4 Sequence Diagram



## 5 References

[1] OpenAI *et al.*, "GPT-4 Technical Report," Mar. 2023, Accessed: Jun. 10, 2025. [Online]. Available: https://arxiv.org/pdf/2303.08774

[2] H. Liu, C. Li, Q. Wu, and Y. J. Lee, "Visual Instruction Tuning," *Adv Neural Inf Process Syst*, vol. 36, Apr. 2023, Accessed: Jun. 10, 2025. [Online]. Available: https://arxiv.org/pdf/2304.08485

[3] T. H. Trinh, Y. Wu, Q. V. Le, H. He, and T. Luong, "Solving olympiad geometry without human demonstrations," *Nature*, vol. 625, no. 7995, pp. 476–482, Jan. 2024, doi: 10.1038/S41586-023-06747-5.

# Use Case 7: AI Enables Intelligent Manufacturing of the Automobile Industry

**Organization:** Huawei Technologies Co., Ltd. & Guangzhou Automobile Group Co., Ltd.

**Country**: China

**Contact Person(s)**:

Zhang Yi, lisaa.zhang@huawei.com

Kong Fanjian, kongfj@aion.com.cn

## 1 Use Case Summary Table

| Item | Details |
|---|---|
| Category | Intelligent manufacturing |
| Problem Addressed | 1. Inspiration for automobile Research and Development (R&D) and design is constrained, leading to prolonged design cycles.<br>2. Failure analysis for full vehicle design involves over 1,000 scenarios, resulting in a substantial workload.<br>3. Critical quality inspection processes in automobile manufacturing rely heavily on manual checks, making outcomes dependent on inspector proficiency.<br>4. Equipment health monitoring relies on manual periodic inspections, causing delays in issue detection and difficulties in accessing elevated or enclosed areas.<br>5. Production scheduling is passively received, causing significant volatility in plans, affecting downstream supply chain preparedness, and increasing inventory costs. |

(continued)

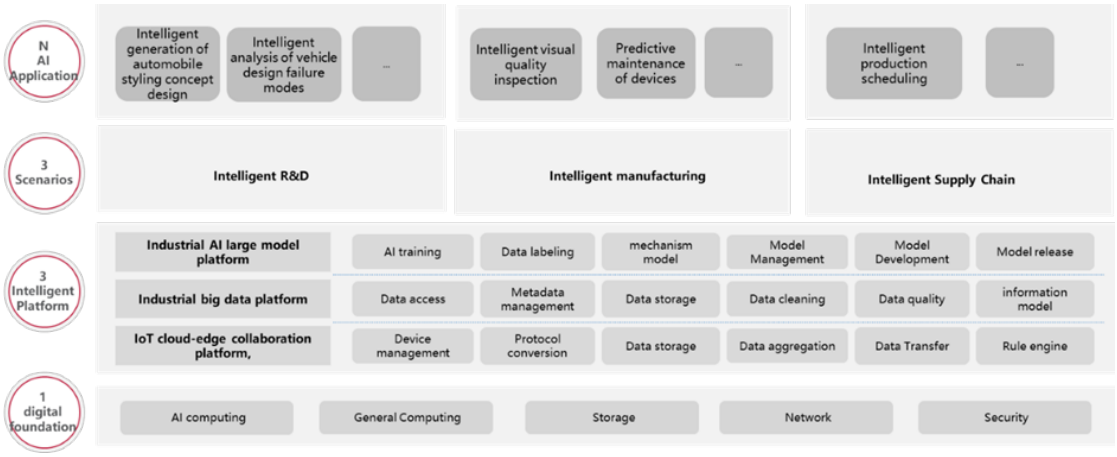
| Item | Details |
|---|---|
| Key Aspects of Solution | 1.Establish a unified architecture platform, we call it "1+3+3+N" architecture. It can enable real-time data interaction between cloud and device terminals, aggregate multi-source/heterogeneous industrial data, provide automobile industry-specific large model development tools, and empower intelligent scenarios across "R&D, production, supply, sales, and service".<br>2. By leveraging Artificial General Intelligence (AGI)-powered text-to-image/ image-to-image generation technologies, the system shortens the concept design phase in styling workflows while reducing product homogenization.<br>3.Apply Large Language Model (LLM) + Graph Retrieval Augmented Generation (RAG) technologies to enhance the efficiency of failure analysis in full vehicle design processes.<br>4.Leverage Computer Vision (CV) large models to minimize repetitive manual defect detection tasks and improve industrial quality inspection efficiency.<br>5.Implement predictive large models and large language models for 24-hour monitoring and analysis of equipment health status, reducing unplanned downtime.<br>6.Using solver technology to achieve intelligent production scheduling, lower inventory costs, and respond to unexpected events proactively. |
| Technology Keywords | Digital foundation, Internet of Things (IoT) cloud-edge collaboration platform, Industrial big data platform, Industrial AI large model platform, AGI, Large language model, Computer vision large model, Solver |
| Data Availability | Private data collected by the Guangzhou Automobile Group (GAC) Co., Ltd., such as equipment predictive maintenance database, intelligent production scheduling dataset, visual quality inspection dataset |
| Metadata (Type of Data) | text, image, video |
| Model Training and Fine-Tuning | Industrial cameras are used to collect image/video from the production line, using the labelled data to fine-tune the CV large model, and use the edge deployment to implement efficient defect detection. |
| Testbeds or Pilot Deployments | The deployment model can be found at [4] |
| Code repositories | Not available |

# 2 Use Case Description

## 2.1 Description

- Traditional automobile manufacturing has a long design cycle. There are more than 1000 scenarios for vehicle design failure analysis, which is a heavy workload. The key quality inspection procedures rely on manual inspection and the ability of quality inspectors. The health status of the equipment is achieved through manual inspection, which has problems such as insufficient timeliness and difficult operations in elevated or enclosed areas. Production scheduling plans are passively received, with great plan fluctuation, affecting the spare parts preparation of the downstream supply chain.

- Solution: GAC group has built a "1+3+3+N" unified architecture system, that is, one digital foundation, three intelligent platforms (IoT cloud-edge collaboration platform, industrial big data platform, and industrial AI large model platform), three scenarios (intelligent R&D, intelligent manufacturing, and intelligent supply chain) and N AI-based applications for automobile manufacturing. The scenarios are described in detail below.




1. Intelligent generation of automobile styling concept design: Using AGI's text-to-image/ image-to-image technology to generate automobile styling concept design, , aiming to address the issues of long concept design cycles and insufficient design creativity in traditional design processes, thereby reducing product homogenization and shortening project development cycles.
2. Intelligent analysis of vehicle design failure modes: Using "LLM + GraphRAG" to understand the vehicle structure and function requirements, automatically generate failure modes and preventive measures, and improve the efficiency of vehicle design failure analysis. RAG includes vehicle body component information, function and failure mode correlation data, internally accumulated design guidelines, historical quality issues, research papers, technical specifications, design validation plans, etc. A function failure analysis, measures, and follow-up implementation that originally required 6 person-days can be shortened to 2 person-days.
3. Intelligent visual quality inspection: Using CV large model to realize the appearance defect inspection of stamped sheet metal parts, paint surface inspection, final assembly specification inspection, etc., to reduce low-value and repetitive manual operations, reduce labor risk and intensity of employees.
4. Predictive maintenance of devices: Using a predictive large model and LLM, 24-hour monitoring and analysis of equipment health status can be achieved. Building a maintenance knowledge base to shorten personnel training cycles and improve equipment availability.
5. Intelligent production scheduling: Using solvers and multi-objective optimization heuristic algorithms are used to realize automatic production scheduling, reduce inventory, and respond to unexpected events. The production capacity will be improved by 10% and the production scheduling time will be shortened by 60%.

**Use Case Status:** The use case is part of a larger research project.

### Partners

**Huawei:** Provides unified digital infrastructure, R&D tool chains, basic large model, IoT cloud-edge collaboration platform, industrial big data platform, and industrial AI large model platform.

**GAC Group:** Focus on the development of AI application scenarios in the "R&D, production, supply, sales, and service" field, and the establishment of high-quality dataset.

## 2.2 Benefits of the use case

1. Building the equipment knowledge base for equipment maintenance enables employees to quickly master professional knowledge, providing more development opportunities in the technical field, enhancing the professionalism and sense of achievement at work. Using AI to replace operations in hazardous environments can reduce labor risks and intensity. Accurate demand forecasting and intelligent production scheduling reduce planning fluctuations, improve resource utilization efficiency, and enhance corporate profitability.
2. In the "1+3+3+N" architecture, multiple intelligent platforms and AI-based applications provide core infrastructure support for automobile manufacturing, inspiring more innovative ideas and business models, enhancing the efficiency of the entire industry chain collaboration, and strengthening industrial competitiveness.

## 2.3 Future Work

The future work of AI enabling the intelligent manufacturing of the automobile industry includes:

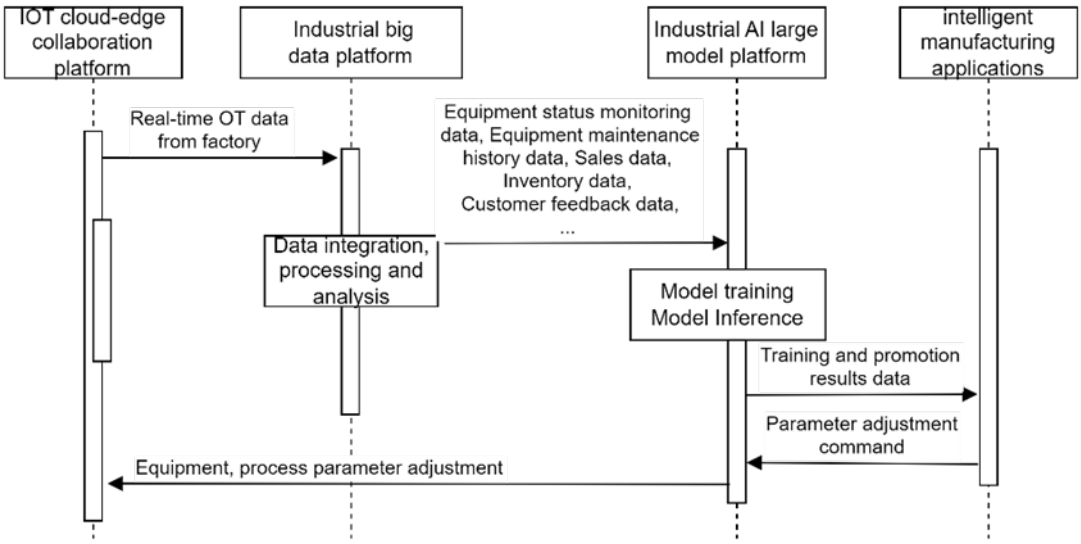
1. Innovative application scenarios: Optimize model development in intelligent R&D, intelligent manufacturing, and intelligent supply chain scenarios, such as intelligent mining tools for intelligent driving scenarios, intelligent vehicle fault diagnosis, and integrated collaboration between production and supply.
2. High-quality dataset: Build high-quality training and evaluation datasets for scenario-based intelligent automobile manufacturing solutions. Datasets are regularly updated to adapt to the evolving automobile intelligent manufacturing technology and market demand changes.
3. Continuous operation: Establish the organizational structure, resource allocation, and operation mechanism for sustainable operation.
4. Industry standards: Develop standards covering the AI model application specifications, data models, data dictionaries, and data security and privacy protection standards in automobile R&D and production, laying a foundation for the healthy and orderly development of the industry. Data models and data dictionaries standards will be developed in ITU-T SG20, while automobile industry AI model application specifications will be developed in ITU-T SG21.

# 3 Use Case Requirements

- **REQ-01**: It is critical to build a unified system architecture which is based on one digital foundation and intelligent platforms to support model development and multiple AI applications in various scenarios such as intelligent vehicle R&D, production, and supply chain.
- **REQ-02**: In production scenario, it is critical that the CV large model is deployed to improve quality detection, efficiency and accuracy, and to reduce labour intensity and risks.
- **REQ-03**: In R&D scenario, it is critical to use a diffusion model for text-to-image/image-to-image generation to accelerate automobile styling concept design. Knowledge graphs, LLM and RAG are used to improve the efficiency of vehicle design failure analysis.

- **REQ-04**: In supply chain scenarios, it is critical to use the solver and multi-objective optimization algorithm to implement dynamic production scheduling, shorten production scheduling time, and reduce inventory costs.

## 4 Sequence Diagram




The IoT cloud-edge collaboration platform collects real-time Operational Technology (OT) data from production devices and pushes the data to the industrial big data platform for data integration, data processing, and data analysis.

Perform model training and inference based on such as historical status data, sales data, inventory data, and customer feedback data.

The training results will be sent to the intelligent application system.

The application system will issue the adjustment instruction according to the training result.

## 5 References

[1] Huawei Technologies Co., "Production digital platform solutions," [Online]. Available: https://e.huawei.com/cn/industries/manufacturing/production-and-supply/produce-digital-platform (accessed: Jun. 15, 2025).

[2] Huawei Technologies Co., "Industrial AI Quality Inspection Solutions," [Online]. Available: https://e.huawei.com/cn/industries/manufacturing/production-and-supply/ai-quality-inspection (accessed: Jun. 15, 2025).

[3] Huawei Technologies Co., "Manufacturing R&D Tool Chain Solutions," [Online]. Available: https://e.huawei.com/cn/industries/manufacturing/innovation/software-rd-toolchain (accessed: Jun. 15, 2025).

[4] Huawei Technologies Co., "HIS Mall," [Online]. Available: https://his.huawei.com/portal/#/sec/intro?servicealias=eai&service_code=hrn:his:servicemarket::service:eai (accessed: Jun. 15, 2025).

# Use Case 8: The Artificial Intelligence Robot for Cement Testing

**Organization**: Dubai Municipality-Dubai Central laboratory

**Country**: United Arab Emirates

**Contact Person(s)**:

Hend Ismail Al Rahma, hirahma@dm.gov.ae

Adnan Mahmoud Al Marzouqi, adnanmm@dm.gov.ae

## 1 Use Case Summary Table

| Item | The Artificial Intelligence Robot for Cement Testing |
|---|---|
| Category | Manufacturing: AI-driven predictive maintenance, process optimization, supply chain management, and robotics<br>AI in Quality Assurance & Material Testing |
| Problem Addressed | Traditional XRF (X-Ray Fluorescence) analysis has long been characterized by its labour-intensive and time-consuming processes. With human involvement in critical steps causing potential errors and increasing safety for laboratory personnel. |
| Key Aspects of Solution | Robotic XRF cement analysers are the first of their kind in the world, marking a significant advancement in the field of materials analysis.<br>The entire process—sample preparation, analysis, and transportation—is fully automated, drastically reducing human involvement in critical steps.<br>This automation minimizes the risk of errors and enhances the safety of laboratory personnel by limiting their exposure to hazardous materials.<br>The AI-driven system not only improves the repeatability and accuracy of analytical results but also dynamically optimizes the testing process.<br>It employs SuperQ software, which works in tandem with the AI to ensure compliance with industry-standard testing methods, including BS EN 196-2: Clause 5 and BS ISO 29581-2:2010. |
| Technology Keywords | XRF analysers, SuperQ software[1], Sam Track (Sample Tracking and Control System), BS EN 196-2, BS ISO 29581-2:2010 |
| Data Availability | Private (Dubai Central Laboratory Portal) |
| Metadata (Type of Data) | Text and visual data<br>Data holds complete path the samples go from start-to finish and is generated from the central quality table |

(continued)

| Item | The Artificial Intelligence Robot for Cement Testing |
|---|---|
| Model Training and Fine-Tuning | The model architecture and a significant portion of the training were conducted at the manufacturer's facility under close monitoring.<br>Fine-tuning, training, and validation, including testing 60 dummy samples manually and then with the automated system, were conducted at the Dubai Central Laboratory to ensure process alignment.<br>The results were then evaluated according to BS EN 196 standards. |
| Testbeds or Pilot Deployments | The automated process was planned at the Dubai Central Laboratory and implemented at a manufacturing facility. The AI/ML system's suitability and acceptance were regularly monitored. |

# 2 Use Case Description

## 2.1 Description

Traditional XRF analysis has long been characterized by labour-intensive and time-consuming processes, where significant human involvement in critical steps introduces potential errors and raises safety concerns for laboratory personnel. The objective has always been to reduce this human involvement, thereby enhancing safety, improving efficiency, and minimizing errors in cement testing.

The introduction of the robotic XRF cement analyser marks a groundbreaking innovation in materials analysis, being the first of its kind in the region[2]. This intelligent robotic laboratory system automates the entire cement testing process—including sample preparation, analysis, and transportation—drastically reducing human intervention. By limiting direct human interaction with hazardous materials, it significantly enhances safety, while also improving operational efficiency and minimizing human errors.

A key element of this system is its integrated AI platform. This AI-powered platform monitors and controls the entire operation, governing hardware components such as weighing balances, bead preparation units, XRF analysers, and robotic apparatus. It dynamically optimizes the testing process, ensuring high repeatability and accuracy. The system employs SuperQ software, which works in tandem with the AI to guarantee adherence to industry standards like BS EN 196-2: Clause 5 and BS ISO 29581-2:2010. Furthermore, the advanced AI seamlessly merges, validates, and distributes generated results with minimal or no human intervention.

By integrating AI and advanced automation, this technology dramatically lowers operational costs and elevates the precision and reliability of cement analysis. Laboratories can now achieve unmatched consistency, efficiency, and safety, ultimately setting new standards for productivity and accuracy within the industry.

### Use Case Status: Robotic XRF system is Active and Implemented

**Partners**

Malvern Panalytical mobilization[3] and establishment partners

## 2.2 Benefits of the use case

This use case aligns with the use case, as the primary objective of the implementation was to minimize human interaction to ensure their safety and well-being.

It also aligns with the introduction of the robotic XRF cement analyser, representing a pioneering innovation in materials analysis. Being the first of its kind in the region, this technology is set to revolutionize the laboratory testing industry by paving the way for complete automation. This will result in increased productivity and efficiency

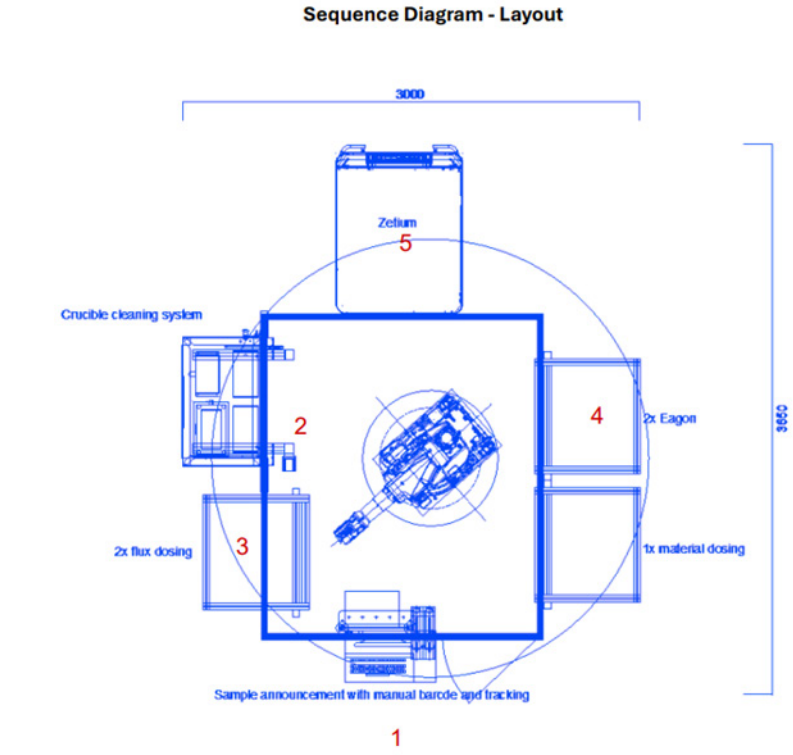
## 2.3 Future Work

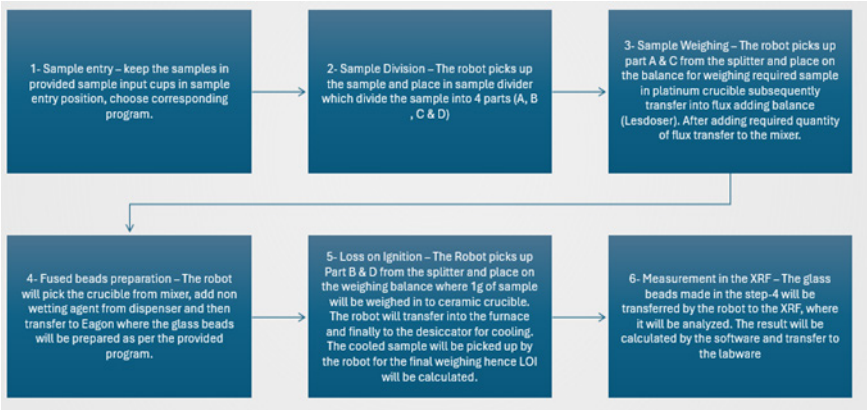
- Planned Automation Expansion
- Future plans include expanding laboratory automation to other labour-intensive test methods like Concrete, Asphalt, and Soil.
- Resource Needs
- The current AI robot for cement testing requires no immediate upgrades. However, future implementations will necessitate test-specific requirements and modifications, which are currently under review.

# 3 Use Case Requirements

- **REQ-01**: It is critical that the AI-driven robotic cement testing system enable seamless interaction and knowledge exchange between robotic hardware components (e.g., robotic arm, XRF analysers, weighing units) and the central data management system for synchronized process execution and monitoring.
- **REQ-02**: It is critical that the AI system optimize its operational knowledge base dynamically to adapt testing parameters based on sample type and class, thus enhancing accuracy, repeatability, and efficiency.
- **REQ-03**: It is critical that the AI system generate comprehensive test reports for consumption by both human analysts and downstream automated systems.
- **REQ-04**: It is critical that the AI-powered laboratory enable validation and secure distribution of testing data across integrated platforms, in compliance with international standards such as BS EN 196-2 and BS ISO 29581-2.
- **REQ-05**: It is critical that the cement testing AI system support modular adaptation to new material types beyond cement and rock, enabling future expansion to include concrete, asphalt, and soil.

## 4 Sequence Diagram




## 5 References

[1] 'SuperQ | XRF Analysis Software | Malvern Panalytical'. Accessed: Jun. 09, 2025. [Online]. Available: https://www.malvernpanalytical.com/en/products/category/software/x-ray-fluorescence-software/superq

[2] 'Dubai Municipality To Test Construction Materials Using AI Robots - Tech Magazine'. Accessed: Jun. 09, 2025. [Online]. Available: https://techmgzn.com/dubai-municipality-to-test-construction-materials-using-ai-robots/

[3] 'Malvern Panalytical | Analytical Instrumentation'. Accessed: Jun. 09, 2025. [Online]. Available: https://www.malvernpanalytical.com/en

# Use Case 9: Synergy of Large Model and Small Model for Port Safety Management

**Organization**: ZTE Corporation

**Country**: China

**Contact Person(s)**:

Liya Yuan, yuan.liya@zte.com.cn

Chen Zhang, zhang.chen9@zte.com.cn

## 1 Use Case Summary Table

| Item | Details |
|---|---|
| Category | Manufacturing: Intelligent transport |
| Problem Addressed | High safety risks in hazardous chemical operations and inefficient manual monitoring at the ports |
| Key Aspects of Solution | Integration of a small edge deep learning model with cloud-based large model in a hybrid architecture for real-time risk detection |
| Technology Keywords | Port, large model, edge-computing |
| Data Availability | Private |
| Metadata (Type of Data) | Visual data |

(continued)

| Item | Details |
|---|---|
| Model Training and Fine-Tuning | Data Collection and Prep: In the early stage, on-site surveillance footage materials were collected, totaling 100GB, including video data of simulated smoke release on-site. After deploying the large model, data screening was conducted, and misidentified data from the scene was added to the model training, achieving dual iteration of both the model and data.<br>Model architecture and Training: As a transformer-based Computer Vision (CV) large model from ZTE, NebulaCV-grounding 2.0 is an innovative open-set object detection model. It not only supports traditional text prompt detection but also introduces visual prompt interactive detection, effectively addressing the challenge of detecting long-tail categories in object detection. The model consists of four core modules: a text prompt encoder, a visual prompt encoder, an image encoder, and an encoder-decoder layer. The text prompt encoder allows users to specify detection categories through text descriptions, while the visual prompt encoder enables users to annotate detection targets interactively on images, significantly enhancing the model's flexibility and practicality. Additionally, Grounding2.0 innovatively incorporates the Mixture of Experts (MOE) architecture in the decoder, significantly improving model performance while maintaining inference speed, achieving a dual optimization of efficiency and accuracy. It is trained on over 2 billion general samples and 60 million industry-specific samples to enhance specialized scenario understanding capabilities.<br>Validation (Performance Metrics) and Fine-tuning: The project employs on-site practical testing to verify accuracy. It utilizes real-world data for Reinforcement Fine-Tuning, optimizing the performance of large models through reward-driven training cycles to enhance their reasoning and generalization capabilities in specific tasks or domains. |
| Testbeds or Pilot Deployments | The solution has been deployed and tested at Taicang Port in Jiangsu, China, with a total of 65 tests conducted.<br>These included 30 smoke videos and 35 non-smoke videos, covering scenarios such as daytime, nighttime, rainy days, and foggy conditions. Based on the deployment test results, the algorithm's accuracy exceeds 90%, with a false alarm rate of less than 5% and a missed detection rate of less than 1%, meeting the requirements of the business scenario. |

# 2 Use Case Description

## 2.1 Description

This case leverages ZTE Nebula visual large model to address the high safety risks in hazardous chemical operations and inefficient manual monitoring at the ports. By integrating a small edge deep learning model with cloud-based Nebula large model in a hybrid architecture[1], the project supports 18 intelligent scenarios spanning four functional domains: personnel management, operational compliance, vehicle tracking, and safety inspections. A unified port safety management platform has been developed to provide **real-time video analytics**, **predictive risk alerts**, and **automated incident forensics**.

Smoke and fire detection is a key scenario in this solution. Traditional edge-based AI camera systems frequently generate false alarms due to environmental complexities. This solution introduces a two-tier AI verification framework:

- The edge small model performs real-time preliminary detection of smoke or flame patterns to trigger immediate alerts
- Cloud-based visual large model conducts secondary validation through contextual analysis, eliminating 92% of false positives.

The system further employs adaptive sampling to prioritize high-risk zones, ensuring comprehensive anomaly coverage without compromising responsiveness.

Since the system's deployment, it has promptly identified multiple safety risks, including workers sitting or leaning against hazardous goods and vehicles illegally crossing container zones. Overall, this has significantly enhanced the port's safety management capabilities.

This project validates the industrial applicability of visual large models while addressing adoption barriers. The hybrid architecture saves cloud computing resources greatly compared to pure large-model solutions, and seamless integration with existing CCTV infrastructure cuts a lot of deployment costs. It also allows rapid onboarding of new risk detection scenarios, ensuring scalability across diverse operational environments.

This approach represents a significant advancement in AI-driven industrial safety monitoring, offering a robust, cost-effective and scalable approach to fire risk management in high-stakes logistics environments.

**Use Case Status:**

Currently, the entire fireworks large model project plan has entered the production stage and is being implemented at multiple ports in China. The plan includes hardware, AI platform, large model algorithms, and applications. The video AI platform is customized for fireworks detection scenarios, with collaborative analysis tasks of large and small models. When users use it, they only need to set template parameters to create a fireworks analysis task, which is then analyzed by the AI platform and provides results to the fireworks warning application.

### Partners

Organization: China Telecom[2]

Role: Provision of cloud platform; collaborative R&D

## 2.2 Benefits of the use case

The project significantly enhances port infrastructure safety management through AI innovation technologies, driving technological advancement in industrial sectors. By reducing implementation costs and increasing system flexibility, it promotes digital transformation of industries and demonstrates the critical role of intelligent technologies in key infrastructure development.

By implementing advanced fire detection technologies, the project strengthens the safety and resilience of logistics infrastructure, effectively mitigating potential production safety risks. This

innovative approach enhances disaster prevention and mitigation capabilities of urban and community environments, protecting vital economic assets and production spaces.

## 2.3 Future Work

### Continuous Improvement

Ongoing data collection and model training will be conducted to ensure the system adapts to more diverse environmental conditions, maintaining the product's cutting-edge capabilities.

### Productization

Based on existing solutions, a standardized product package will be developed, including hardware (e.g., cameras) and software (AI algorithms and cloud deployment solutions). This will enable rapid deployment in other ports or logistics centers, reducing implementation barriers.

Platform Development: A cloud-based service platform will be established, integrating visual large models and other auxiliary tools to provide users with on-demand AI capabilities. The platform will support a multi-tenant model, allowing customized services according to different requirements.

### Collaboration

Explore collaboration opportunities with other industry partners, such as insurance companies, to innovate and develop new business models.

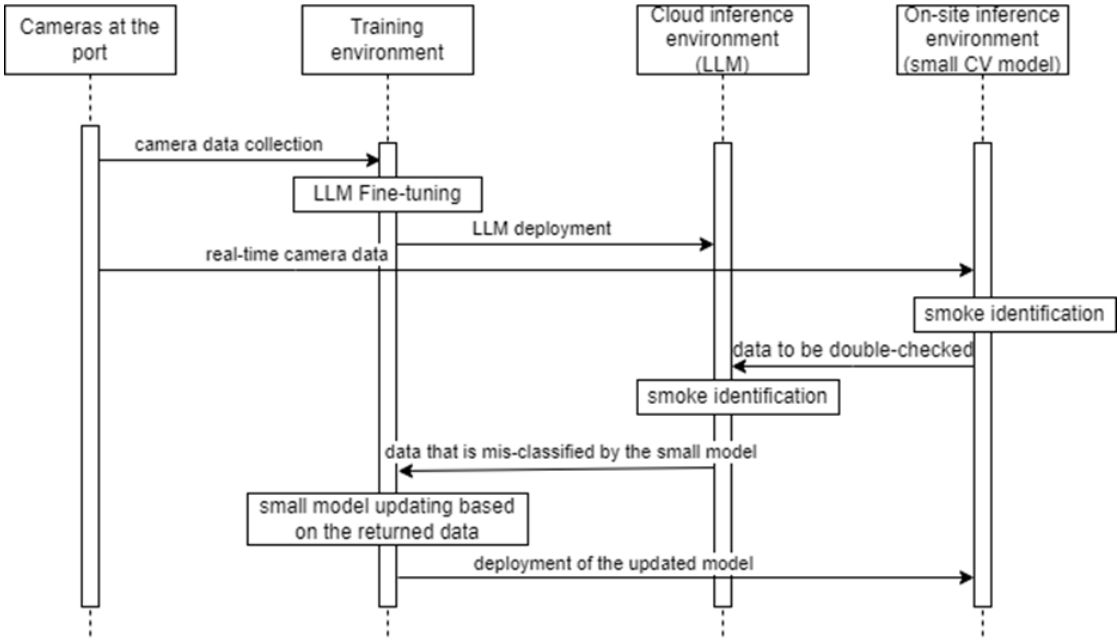
### Standards Development

Actively participate in the development of relevant national and global standards e.g., in SG21, sharing our expertise and best practices in the field of AI surveillance for safety management.

# 3 Use Case Requirements

- **REQ-01**: It is critical that hybrid edge-cloud architecture is required to enable real-time detection (edge) and contextual validation (cloud), minimizing latency and cloud resource usage.
- **REQ-02**: It is of added value that Edge-optimized lightweight models are required to run efficiently on low-power devices while maintaining detection accuracy.
- **REQ-03**: It is critical that fine-tuning for both cloud large model and edge small model is required to improve validation accuracy and adapt to evolving environmental conditions
- **REQ-04**: It is critical that unified management interface is required to centralize monitoring of personnel, vehicles, compliance, and safety inspections across all port zones.
- **REQ-05**: It is expected that dynamic resource allocation is required to balance edge/cloud workloads based on risk severity and operational priorities.
- **REQ-06**: It is critical that data security protocols are required to encrypt video streams and operational data across edge/cloud layers.

## 4 Sequence Diagram




## 5 References

[1] ZTE, "Special topic: Nebula industry big model." Accessed: Jun. 10, 2025. [Online]. Available: https://www.zte.com.cn/china/about/magazine/zte-technologies/20230/10-cn.html

[2] "China Telecom Corporation Limited." Accessed: Jun. 10, 2025. [Online]. Available: http://www.chinatelecom.com.cn/

# Use Case 10: AI-Based eXtended Reality Training for Automation

**Organization**: BSD design

**Country**: Italy

**Contact Person(s):**

Alessandro Pollini

Alessandro Pollini, alessandro.pollini@bsdesign.eu

Angela Di Massa, angela.dimassa@bsdesign.eu

## 1 Use Case Summary Table

| Item | Details |
|---|---|
| Category | Manufacturing |
| Problem Addressed | Industry training for automation, specifically programming a metal bending robotic cell, utilizes a practical approach combining Computer Numerical Control (CNC) and collaborative robot programming with hands-on exercises. This experience-based method accelerates learning compared to traditional methods. |
| Key Aspects of Solution | E-XTENDIT is a Modular XR (Extended Reality) toolkit for enhancing human-centred manufacturing based on (1) a gamified strategy for providing engaging and effective industrial training, (2) the AI personalization to allow automatic personalization training material and adjust contents in VR environments; (3) XR interfaces, interactions and solutions to provide guidance and support user decision-making in real-time operation and training support; (4) preliminary skills assessment and evaluation, through field experiments and evaluation campaigns. |
| Technology Keywords | eXtended Reality |
| Data Availability | Private |
| Metadata (Type of Data) | Visual |

(continued)

| Item | Details |
|---|---|
| Model Training and Fine-Tuning | Personalization involves data analysis and insights generation and it follows two different approaches with the aim of 1) recommending optimal levels of automation and 2) predicting the expected level of productivity, thereby improving operational efficiency and identifying potential areas for improvement. Firstly, a statistical analysis are performed on the integrated data to develop a data-driven rule-based approach, augmented with expert knowledge, resulting in a set of hard-coded rules. Additionally, machine learning techniques are employed to identify unexpected characteristics influencing productivity, utilizing explainable models like decision trees and more complex models such as Support Vector Machines and Artificial Neural Networks. Model-agnostic methods like SHapley Additive exPlanations (SHAP) and Local Interpretable Model-agnostic Explanations (LIME) are used to interpret these complex models. Finally, an easy-to-use Augmented Reality (AR) add-on is developed, leveraging insights from rule-based and machine-learning approaches. This add-on is provided following a containerized microservice approach for seamless integration and deployment. |
| Testbeds or Pilot Deployments | Advanced Manufacturing Technologies (AMT) and the Deep Tech approach, converging digital and physical realms, are crucial. Industrial operators need new emotional and relational skills for interacting with intelligent machines, while training designers require a holistic, multidisciplinary approach. Extended Reality and the Industrial Metaverse (XR/IM) are key to innovating production and fostering a skilled, resilient, and creative workforce, essential for successful industrial transitions and competitiveness. Industry training for automation, specifically programming a CNC machines and collaborative robot with hands-on exercises. This experience-based method accelerates learning compared to traditional methods. |
| Code repositories | Not available |

# 2 Use Case Description

## 2.1 Description

- The project aims to define, support and accelerate the adoption of adaptive distributed XR Human-Machine Interaction (HMI) in the industry, by introducing to the market a modular and distributed XR Toolkit for enhancing human-centered manufacturing based on adaptive, personalized and distributed XR industrial machine interfaces for a safer, more inclusive and well-integrated working environment.
- XTENDIT integrates XR, Internet of Things (IoT), AI and mobile devices for promoting a more adaptive and informed use of machine automation to help industries and machine operators. Specifically, XTENDIT is a XR Toolkit supporting human decision-making and thus well-being, safety and productivity of the workforce in manufacturing settings. As an example of the intelligence augmentation paradigm, that sees human and AI working together as a symbiotic system, XTENDIT proposes AI to augment human users' decision-making process and capabilities by providing otherwise hidden or inaccessible data-driven insights.

XTENDIT use case's objectives are:

- 1_ To implement an interaction design model for Human-Centered Automation and AI-Human Augmentation, based on personalized, adaptive, accessible, distributed and context-sensitive user interfaces;

- 2_ To validate a modular, personalized and distributed XR Toolkit for enhancing human-centered manufacturing;
- 3_ To validate and commercialize the XR Toolkit towards sustainable resource consumption, effective production and empowerment of workers.

XTENDIT represents a transformative advancement in the integration of technologies within the industrial sector, particularly focusing on the development of adaptive, personalized, and distributed XR industrial machine interfaces under the framework of Industry 5.0. XTENDIT integrates a multiplexed technology integration approach that will jointly leverage

- user ergonomics, by using pre-collected human factors data on personal values and attitudes towards automation, including trust, acceptance, reluctance, fear and anxiety that are collected as individual dispositions by mean of an ex-ante self-reported questionnaire on the workers;
- machine Information, by integrating pre-existing detailed information about the current machines;
- contextual Information about the work scenario, including both the user scenario, i.e. whether the user is programming, supervising, performing manual operations, or maintaining equipment; and the productivity, waste, and efficiency data collected through IoT components.

As for workforce resilience improvement, XTENDIT product innovation will be based on a significant collection of EU-representative ergonomics data related to human-machine interaction during testing machine measurement. In particular, XTENDIT worker centricity relies on the investigation of users' psychological (knowledge, skills, and attitude towards automation), and cognitive dimensions (learning, inspection, awareness, decision-making), by means of a self-reported questionnaire. The skills assessment will be provided before, during and after the training with the XR Toolkit in order to measure improvements in participants' skills and in improving productivity but also the user experience and engagement level. According to the knowledge collected during the training needs' analysis, the trainers might develop context intelligent training - trainees will be engaged in on-field AR interactions that sense and respond to the environment, allowing them to exercise with relevant situations that may occur during production.

On this basis, XTENDIT elaborates a novel XR and MR (extended and mixed reality) interface design paradigm to accomplish the empowerment of workers and the optimization of production by employing information provision and dynamic support to decision-making in an effective and timely manner, according to the operator's physical and cognitive ergonomics.

At the methodological level, XTENDIT implements a unique interaction design-led method for resilient and human-centered Advanced Manufacturing, based on mixed materials and prototypes, and on the application of a new, innovative, and multidisciplinary method for product innovation, also including human body interaction, with the possibility of integrating wearable technologies to adapt the interfaces on the basis of situational data, e.g. the user location.

At the organizational level, XTENDIT represents a case of the modernization of equipment and protocols for research, assessment, and modeling industrial production operations, including machine interfaces' adaptations that are meaningfully rich for the end users.

XTENDIT significantly reduces the burden of discomfort, and, consequently, the delays, the blocks and the inefficiencies of a not integrated human-automation organization. The product

will be enriched by implementation guidelines to support replicability and further exploitation of results in terms of design of intelligent software algorithms and advanced human-machine interfaces.

In particular, XTENDIT might apply to all the different place-based training scenario, e.g. on-the-job physical experiences, that offer the possibility to both scan and create the 3D model of a system (e.g. machine, vehicle, industrial line), and position markers to implement points of interests around which is possible to create the interactive XR training. Replicability is thus assured by the XR training authoring tool that makes the 3D contents both available by means of the adoption of the AR and VR apps.

### Partners

Alma Mater University of Bologna, with the involvement of the Department of Architecture, through the Advanced Design Unit (ADU), and the Department of Industrial Engineering collaborated on the project. The UNIBO Advanced Design Unit (ADU [14]) operates in the Department of Architecture through the tracks on Industrial Design and Advanced Product and Service Design, which have been established since 2013 to train students on design-driven approaches to innovation. The Department of Industrial Engineering of the University of Bologna has been working for years on issues related to human-machine interaction and augmented and virtual reality systems as new typologies of user interfaces. The UNIBO team has contributed to the design of novel interaction paradigms and adaptive interfaces, and to the iterative and incremental implementation of XTENDIT field experiments.

## 2.2 Benefits of the use case

1. By considerably improving workers upskilling and capability of managing automated machines, XTENDIT promote operators that are resilient and creative problem solvers, and that impacts on improvement of number of components produced per unit of time after scaling up, reduction in production downtime due to technology adoption and integration, number of error reduction and number or rework, reduction of onboarding time for new workers;
2. By enhancing industry training capability, XTENDIT aims to increase the technological literacy rate in the target population, the upskilled employees rate (pre and post skills assessments), and the rate of talents attracted by the industry.

## 2.3 Future Work

XTENDIT aims to extend its exploitation of adaptive distributed XR HMI, by implementing and introducing to the market a module for XR gamification strategies for Human-Centered Augmentation. Based on personalised, gamified, adaptive, distributed and context-sensitive user interfaces, XTENDIT will include gamification in AI-driven personalized strategies that adapts dynamically to user needs, engagement and experience levels, by means of collecting and analyzing human-machine interaction data to continuously optimize the training experience and interface personalization.

The ultimate scope is to enhance the work and training environment to make manufacturing jobs more safe and attractive for young generations. The accomplishment of this result will be informed by the assessment of the impact of training scenarios in XR environments on problem-solving capabilities within industrial settings, through gamified strategies, such as challenges,

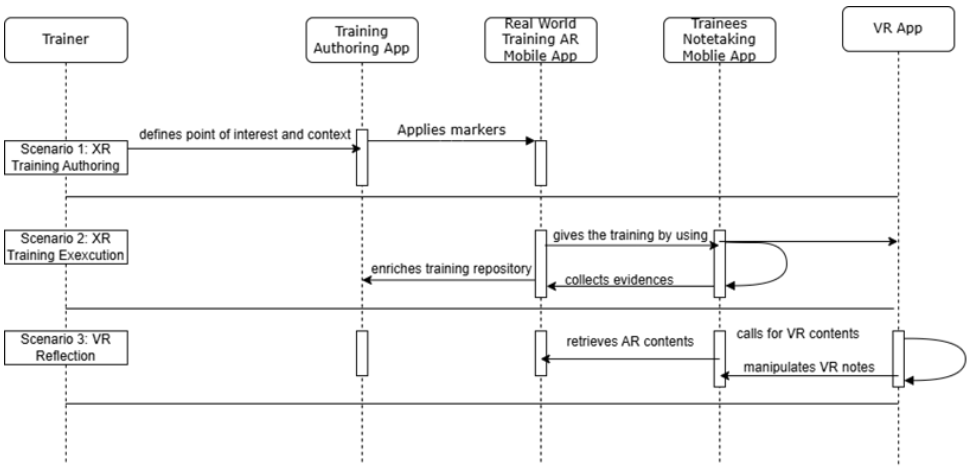
score system and informational hints, by measuring time-to-resolution, training effectiveness, engagement, and skills acquisition, supported by data-driven insights from field experiments and user evaluations.

## 3 Use Case Requirements

- **REQ-01**: It is critical that training be accessible in a multilingual approach, given language barriers and time zone differences in remote training,
- **REQ-02**: It is critical that trainees personalize according to the varying levels of prior knowledge and skills.
- **REQ-03**: It is critical that Augmented reality (AR) training be provided in different use cases, such as: distance learning, remote maintenance and safety instructions.
- **REQ-04**: It is critical that the system addresses related challenges, such as AR hardware and software costs that can be expensive and the technical skills required to create and use AR content.
- **REQ-05:** Customised On-the-Job training provided by experienced operators. This approach allows for quickly learning the specific knowledge of each machine and receiving advice and tips not available in standard courses.
- **REQ-06**: Safety: training must include a section on safety to prevent accidents during machine use;
- **REQ-07**: Operators' limited knowledge of the system or machine, technical issues with the software or the machine itself, training expert technicians is costly (both in terms of time and economically).
- **REQ-08**: Lack of adequate contextualised guides during training. Trainers must have appropriate materials accompanying the training, such as user guides and interactive on-machine tutorials.
- **REQ-09**: Need to diversify activities to reduce physical fatigue. Operators express concern about physical fatigue resulting from standing for long periods in front of the machine. To alleviate this issue, it is suggested to diversify their activities by integrating complementary ones that can be performed nearby the machine, such as training short tasks.
- **REQ-10**: The system must support on-the-job training sessions with real machinery and equipment. These sessions can range from 1-2 hours to a full day.
- **REQ-11**: The system must provide limited dependence on automation. Automation does not completely replace human skills. Operators are essential for supervision, maintenance, and solving complex problems.
- **REQ-12**: Main training sessions topics: monitoring industrial machines and production processes; execution of production and assembly procedures.
- **REQ-13**: Need for standardised training providing an overview of the system and program hierarchies (such training would enable operators to move more autonomously across different machines);
- **REQ-14**: Need to address scalability and replicability on a large scale.
- **REQ-15**: Machine or software operations: operators must learn to use the machine or the various functions of the software to program the machines (e.g.: CNC);
- **REQ-16**: Maintenance: operators need to acquire basic skills for routine machine maintenance or troubleshooting when needed. Often, operators turn to the trainer outside of training hours for on-the-job support through calls or video monitoring (e.g.: Team Viewer).

- **REQ-17**: Focus on practical training. Training primarily focuses on the practical use of machines or software. Generally, support materials such as slides or manuals are not used, except for asynchronous online training (e.g.: video courses). Participants typically try to memorise what they see through practice and repetition of the steps shown by the trainer, but without the support of traditional technologies or tools (e.g.: written notes).
- **REQ-18**: Need for personalised training. Training should be tailored to the different needs and levels of experience of operators. Sometimes, this turns out to be the main critical issue encountered.
- **REQ-19**: The system must convey a positive perception of automation. Trainers generally view automation as a way to simplify processes, increase precision, and reduce human error. Trust in automation is present, but not absolute. Operators have more trust in machines when there is adequate support or when they understand how the system works. Safety is a key factor for trust in automation.

## 4 Sequence Diagram




skills acquisition, supported by data-driven insights from field experiments and user evaluations.

## 5 References

[1] Bologna, J.K., Garcia, C.A., Ortiz, A., Ayala, P.X., Garcia, M.V: An augmented reality platform for training in the industrial context. IFAC-PapersOnLine 53, 197–202 (2020)

[2] Caballini, C., Carboni, A., Boero, F., Parodi, F., Valentini, I., Paolucci, M., ... & Pagano, S. (2023). Augmented reality and portable devices to increase safety in container terminals: the testing of A4S project in the port of Genoa. Transportation research procedia, 69, 344-351.

[3] Colì, E., Paciello, M., Falcone, R., Saleri, G., Pepe, M., & Pedon, A. (2019). Interpersonal trust in adolescence: a preliminary study on online/offline social interactions and life satisfaction. International Journal of Developmental and Educational Psychology, INFAD Revista de Psicología, N°2, 2019. ISSN: 0214-9877. pp:343-352.

[4] Dodoo, J. E., Al-Samarraie, H., Alzahrani, A. I., & Tang, T. (2025). XR and Workers' Safety in High-Risk Industries: A Comprehensive Review. Safety Science, 185, 106804.

[5] Endsley, M. R. (2018). Automation and situation awareness. In Automation and human performance (pp. 163-181). CRC Press.

[6] Grandi, F., Prati, E., Mangia, G., & Peruzzini, M. (2023, June). Development of an AR-Based Application for Training of Warehouse Operators. In International Symposium on Industrial Engineering and Automation (pp. 121-132). Cham: Springer Nature Switzerland.

[7] Leng, J., Sha, W., Wang, B., Zheng, P., Zhuang, C., Liu, Q., ... & Wang, L. (2022). Industry 5.0: Prospect and retrospect. Journal of Manufacturing Systems, 65, 279-295.

[8] Kiourtis, A., Mavrogiorgou, A., Makridis, G., Kyriazis, D., Soldatos, J., Fatouros, G., ... & Falsetta, M. (2024, October). XR5. 0: Human-Centric AI-Enabled Extended Reality Applications for Industry 5.0. In 2024 36th Conference of Open Innovations Association (FRUCT) (pp. 314-323). IEEE.

[9] Mourtzis, D., Angelopoulos, J., & Panopoulos, N. (2022). Operator 5.0: A survey on enabling technologies and a framework for digital manufacturing based on extended reality. Journal of Machine Engineering, 22.

[10] Onnasch, L., & Hösterey, S. (2019, November). Stages of decision automation: Impact on operators' role, awareness and monitoring. In Proceedings of the Human Factors and Ergonomics Society Annual Meeting (Vol. 63, No. 1, pp. 282-286). Sage CA: Los Angeles, CA: SAGE Publications.

[11] Pierdicca, R., Frontoni, E., Pollini, R., Trani, M., & Verdini, L. (2017). The use of augmented reality glasses for the application in industry 4.0. In Augmented Reality, Virtual Reality, and Computer Graphics: 4th International Conference, AVR 2017, Ugento, Italy, June 12-15, 2017, Proceedings, Part I 4 (pp. 389-401). Springer International Publishing.

[12] Romero, D., Stahre, J. (2021) 'Towards The Resilient Operator 5.0: The Future of Work in Smart Resilient Manufacturing Systems'. Procedia CIRPw

[13] Xu, X., Lu, Y., Vogel-Heuser, B., & Wang, L. (2021). Industry 4.0 and Industry 5.0—Inception, conception and perception. Journal of Manufacturing Systems, 61, 530–535. https://doi.org/10.1016/j.jmsy.2021.10.006

[14] Universita Di Bologna, "Advanced Design Unit," [Online]. Available: http://adu.unibo.it (accessed: Jun. 15, 2025).

# Use Case 11: Industrial AI Quality Inspection Based on IndustryGPT Multimodal Industrial Large Model

**Organization**: Shenzhen Smartmore Co.Ltd

**Country**: China

**Contact Person**:

Chi Zhang, chi.zhang@smartmore.com

## 1 Use Case Summary Table

| Item | Details |
|---|---|
| Category | Manufacturing: Intelligent inspection, smart manufacturing, industrial large model application |
| Problem Addressed | Traditional industrial quality inspection methods are plagued by inefficient manual processes, cumbersome data annotation and cleaning, and slow model training. These challenges hinder productivity and delay the adoption of AI solutions, making it difficult for industries to achieve rapid implementation and application of intelligent quality control. |
| Key Aspects of Solution | IndustryGPT, a large industrial model, revolutionizes industrial AI by solving inefficient manual inspection and complex data challenges. It enables "development through conversation" for rapid AI implementation.<br>It operates via the SMore ViMo intelligent cloud platform, an end-to-end system for industrial intelligence[1]. IndustryGPT comprehends ML processes, using natural language to drive ViMo recommendations and automate development steps.<br>On the production line, IndustryGPT transforms the industrial quality inspection all-in-one machine, replacing manual checks. This system precisely detects and screens products with significantly superior speed and accuracy. |
| Technology Keywords | IndustryGPT industrial large model, intelligent inspection, smart manufacturing, AIGC |
| Data Availability | Data is private |
| Metadata (Type of Data) | Industrial images (over 3 million), equipment operation data, sensor data (visual, thermal imaging, gas concentration, vibration, sound), text data (over 50 billion tokens), defect annotation data, multi-industry knowledge bases (five disciplines: optics/mechanics/electronics/computing/software, eight industries).<br>In the application scenarios of IndustryGPT, the data type is primarily defect image data. |

(continued)

| Item | Details |
|---|---|
| Model Training and Fine-Tuning | IndustryGPT integrates deeply with the ViMo cloud platform, acting as an intelligent assistant for industrial AI model training. Users can now prepare data, select models, and configure parameters using natural language, even voice commands.<br>A built-in industrial knowledge base ensures the system accurately understands complex requirements like "scratch detection." It provides full-process intelligent support, from recommending data augmentation to dynamically optimizing hyperparameters.<br>This dramatically shortens development cycles, turning days into hours. It boosts new user efficiency by five times and lets algorithm engineers manage triple the projects, fostering agile AI deployment and continuous quality improvement. |
| Testbeds or Pilot Deployments | This technology has already formed mature application cases in international leading enterprises such as Carl Zeiss, Airbus, Bosch, Canon, Continental Group, and Schaeffler |

# 2 Use Case Description

## 2.1 Description

IndustryGPT, an industrial large model, aims to significantly boost industrial productivity. It tackles key inefficiencies in current industrial AI quality inspection: the low efficiency of traditional manual quality inspection, the cumbersome nature of data annotation and cleaning, and the slow pace of model training.

IndustryGPT achieves this by revolutionizing industrial software operation and interaction. It deploys software functions through its interaction with the SMore ViMo intelligent cloud platform. The ViMo platform is a comprehensive industrial intelligence platform that manages the entire process from model creation to application deployment, drawing on experience from over 1000 specialized industry scenarios.

By leveraging its understanding of the ViMo intelligent cloud platform's technical documentation, IndustryGPT can comprehend the machine learning model development process. When users issue natural language commands, IndustryGPT initiates ViMo's intelligent recommendation system. This system then automatically generates algorithm solutions and handles tasks such as data annotation, parameter tuning, and model training.

Integrating IndustryGPT with the ViMo intelligent cloud platform directly on the production line transforms it into the "brain" of the industrial quality inspection integrated machine. This machine is embedded into the production line to replace traditional manual inspection, enabling the detection and screening of faulty products.

In a specific classification detection example, the solution chosen by IndustryGPT achieved remarkable results: it took only 5 minutes and 27 seconds with 100% accuracy and recall rates. In contrast, manual training for the same detection took significantly longer, at 9 minutes and 31 seconds, and yielded much lower performance with only a 50% accuracy rate and 66.7% recall rate.

### Use Case Status: Production line application

**Partners**

This technology is a solution developed entirely independently by SmartMore Technology.

This innovative technology was developed solely by **SmartMore Technology**. Their core expertise lies in **visual AI**, which forms the backbone of their holistic industrial intelligent quality inspection solution. This integrates hardware, software, data, and deployment, delivering high-precision defect detection, dimensional accuracy, and assembly quality checks. SmartMore's solution is designed for rapid integration into production lines, boasting a ≥99% defect detection rate and ≤5% false positive rate, ultimately reducing costs and boosting productivity.

## 2.2 Benefits of the use case

1) the use case enables rapid implementation and application of intelligent quality control
2) the use case solves the cumbersome nature of data annotation and cleaning, and the slow pace of model training.
3) it generates algorithm solutions and handles tasks such as data annotation, parameter tuning, and model training
4) the use case enhances manufacturing process automation and inspection accuracy, supports flexible production, optimizes resource utilization, pushes industries toward more efficient and sustainable digital transformation.

## 2.3 Future Work

The development direction of the industrial large model IndustryGPT mainly focuses on the expansion of intelligent manufacturing industry scenarios, the refinement and commercialization of customized solutions, human-machine collaboration, cross-industry applications, and sustainable development.

IndustryGPT will focus on providing intelligent solutions to address issues such as flexible production equipment and rapid line changes.

IndustryGPT will empower manufacturers to quickly achieve model training, software integration, and production line deployment even without professional algorithm engineers, greatly improving production efficiency and product quality.

IndustryGPT will also provide personalized solutions based on the needs of different industries, serving as an intelligent assistant to enhance human-machine collaboration, improving safety and efficiency.

In addition, it will explore the potential in multiple fields such as energy, logistics, and emergency response to drive digital transformation.

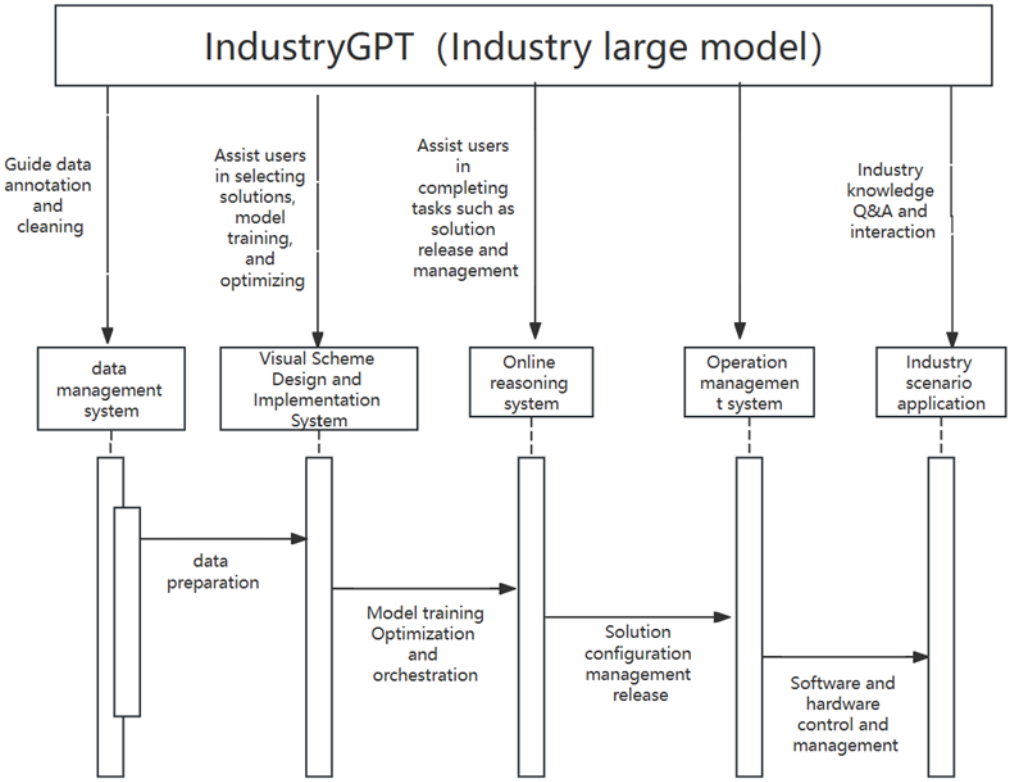
At the same time, IndustryGPT is committed to optimizing resource allocation, improving business quality, supporting enterprises in achieving sustainable development goals, and becoming an important force in promoting industrial intelligence

## 3 Use Case Requirements

The technology should meet the following technical requirements in practical applications.

- **REQ-01**: It is critical that the solution achieves a defect detection accuracy rate of at least 85%.
- **REQ-02**: It is critical that the precision, recall, and F1 scores for defect detection are not lower than 80%.
- **REQ-03**: It is critical that the system can automatically generate algorithm solutions and manage the entire machine learning workflow (data annotation, parameter tuning, training) from natural language commands.
- **REQ-04**: It is critical that the solution can replace traditional manual inspection by being embedded into the production line to detect and screen problematic products.
- **REQ-05**: It is critical that the end-to-end development cycle for new AI models is significantly shortened, reducing time from days to hours, to enable rapid industrial AI implementation.

## 4 Sequence Diagram




## 5 References

[1] "SmartMore - SMore ViMo." Accessed: Jun. 09, 2025. [Online]. Available: https://global.smartmore.com/smore-vimo.html

# Use Case 12: AI-Based System for Assessing Health and Moisture Content of Concrete

**Organization:** Mbeya University of Science and Technology

**Country**: Tanzania

**Contact Person:**

Godfrey Samwel Butete, godsamwel4@gmail.com, +255757315443

## 1 Use Case Summary Table

| Item | Details |
|---|---|
| Category | Manufacturing |
| Problem Addressed | Manual inspection of concrete structures is time-consuming, often destructive, and lacks real-time monitoring capability. |
| Key Aspects of Solution | AI-based non-destructive evaluation using Convolutional Neural Networks (CNNs) for crack detection and sensor data for moisture analysis. |
| Technology Keywords | AI, CNN, Regression, IoT Sensors, Edge Computing, Structural Health Monitoring |
| Data Availability | Combination of public datasets (SDNET2018 dataset [3]) and private data collected at MUST labs. |
| Metadata (Type of Data) | Visual data (images of cracks), sensor data (moisture, strain, load) |
| Model Training and Fine-Tuning | CNNs trained with image datasets; regression models trained on sensor logs; uses transfer learning. |
| Testbeds or Pilot Deployments | University lab environment, local construction firms (planned) |
| Code repositories | Not yet hosted in Github |

# 2 Use Case Description

## 2.1 Description

### Problem Statement

Manual inspection of concrete structures is labor-intensive, time-consuming, and prone to inaccuracies. Traditional methods often involve destructive testing, delayed reporting, and reactive maintenance, which compromises structural safety and increases operational costs.

### Objectives

*Main objective*

Develop an AI-powered, non-destructive system to assess concrete health and moisture in real-time.

*Specific objective*

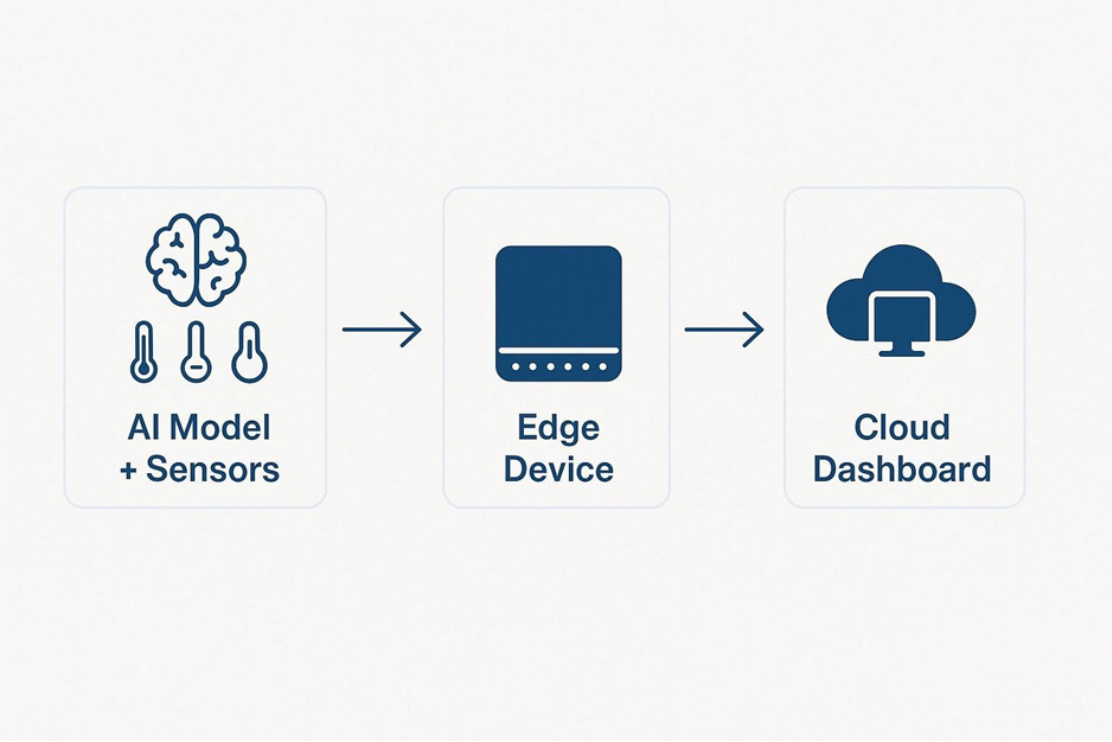
Using sensor array subsystem for obtaining signal

Edge Computing for integrating both Preprocessing and Health Index Calculation.

Also using AI inference for Crack Detection and Moisture Regression.

Deliver actionable insights via cloud dashboards accessible through mobile platforms.

The diagram below illustrates the working of the use case



### Significance

This system aims to revolutionize infrastructure maintenance by integrating AI with civil engineering tools, enabling predictive decision-making, extending the lifespan of critical infrastructure, and aligning with sustainability and smart city goals.

This use case aims to address the inefficiencies in traditional concrete monitoring methods by introducing an AI-powered system capable of real-time, non-destructive evaluation of concrete structures. The system uses CNN models to detect cracks and regression-based models to predict moisture content using sensor input. It helps construction engineers and maintenance teams identify early-stage damage and moisture anomalies, thereby enhancing infrastructure durability, safety, and resource optimization.

**Use Case Status**: Concept validated, data collection and model development ongoing.

### Partners

Mbeya University of Science and Technology (research and lab testing), local construction firms (validation and implementation), sensor vendors (hardware support).

## 2.2 Benefits of Use Case

The project contributes to industry, innovation, and infrastructure, sustainable cities and communities, and responsible consumption and production by improving infrastructure resilience, supporting early detection of construction issues, and reducing material waste.

## 2.3 Future Work

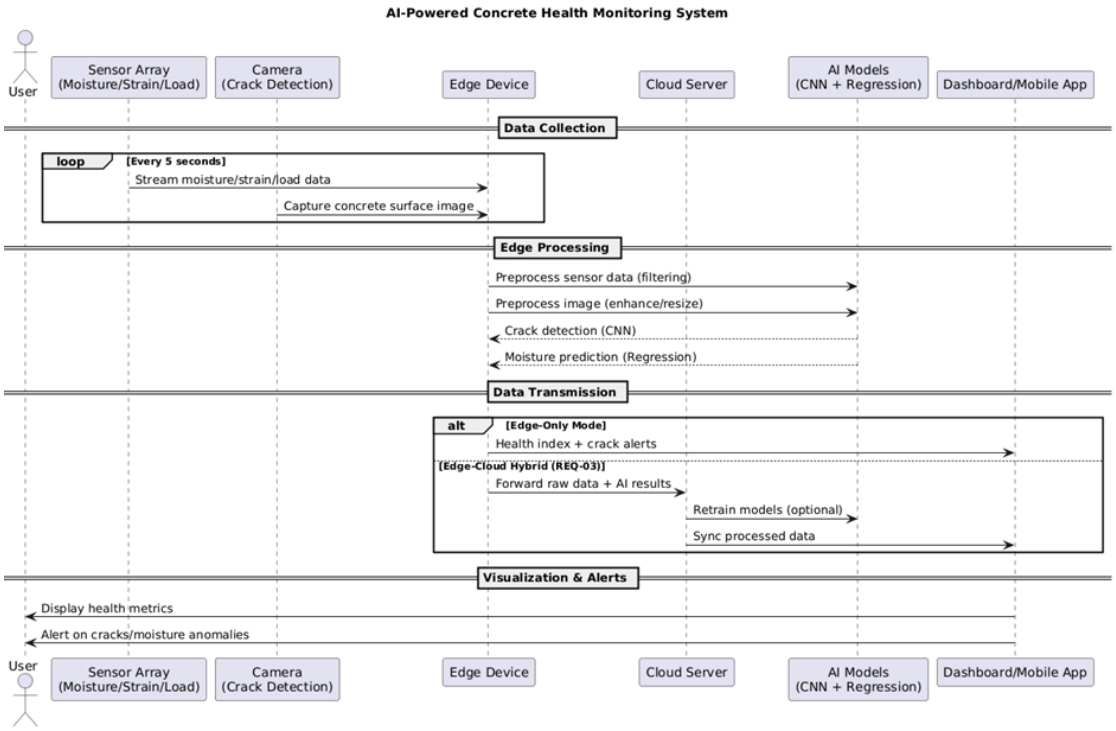
Planned steps include:

- Expanding dataset using field data
- Collaborating with construction firms for real-world pilots
- Improving model accuracy and generalization
- Mobile application development for on-site use
- Integration with Building Information Modeling (BIM) platforms

# 3 Use Case Requirements

- **REQ-01**: Integration of sensors (moisture, strain, load) is critical for data collection
- **REQ-02**: Real-time crack detection using camera images is critical
- **REQ-03**: Data transmission to an edge or cloud server is critical
- **REQ-04**: AI model deployment on edge device or cloud platform is critical
- **REQ-05**: User interface/dashboard for visualization and alerts s critical

## 4 Sequence Diagram




## 5 References

[1] Cha, Y.J., Choi, W., & Büyüköztürk, O. (2017). Deep learning-based crack damage detection using convolutional neural networks. Computer-Aided Civil and Infrastructure Engineering.

[2] Gopalakrishnan, K., et al. (2020). Machine learning for civil infrastructure: a perspective. IEEE Access.

[3] Mendeley Data, "Concrete Crack Images for Classification (SDNET2018)," [Online]. Available: https://data.mendeley.com/datasets/5y9wdsg2zt/2 (accessed: Jun. 15, 2025).

[4] UN SDG Official Report (https://sdgs.un.org/goals)

# Use Case 13: AI Control System for Fusion Energy

**Organization**: Next Step Fusion

**Country**: Luxembourg

**Contact Person(s)**:

Aleksei Zolotarev, info@nextfusion.org

Georgy Subbotin, gs@nextfusion.org

Olga Spiridonova, os@nextfusion.org

## 1 Use Case Summary Table

| Item | Details |
|---|---|
| Category | Manufacturing |
| Problem Addressed | Developing robust and reliable plasma control systems is one of the challenges in the transition from experimental fusion devices to fusion power plants. Plasma control is crucial for maintaining the high temperatures and magnetic confinement necessary for sustained fusion reactions. An effective plasma control system must coordinate numerous magnetic coils and adjust their voltages thousands of times per second to keep the plasma stable - tasks current systems can't reliably manage, highlighting the need for AI-powered solutions. |
| Key Aspects of Solution | • **Next Step Fusion Simulator (NSFsim) -** our in-house physics-based tokamak simulator used to model plasma behavior and train RL agents safely and efficiently. More information on our NSFsim can be found at [1]<br>• **Machine Learning model** that performs real-time control of plasma shape and position by processing raw magnetic diagnostics data and outputting commands to adjust currents in the tokamak's magnetic coils.<br>• **Reinforcement Learning algorithms** to train ML models for real-time plasma parameter control, as well as for training an ML model for plasma shape reconstruction. |
| Technology Keywords | Machine Learning, Reinforcement Learning, Plasma Control System, Plasma Simulation |
| Data Availability | Private (Data can be reached through the DIII-D national user facility program. More information about joining the program can be found at [2]) |
| Metadata (Type of Data) | • Time-series experimental data from fusion devices, including magnetic sensor measurements (e.g., magnetic probes, flux loops)<br>• Control command for power supply (converted to voltage commands by power supply hardware)<br>• Additional synthetic data is generated using our NSFsim tokamak simulator to supplement real-world data and accelerate model development. |

(continued)

| Item | Details |
|---|---|
| Model Training and Fine-Tuning | Effective shape and position control via magnetic system is provided with the use of Reinforcement Learning algorithms, mainly Soft Actor-Critic. Our approach leverages privileged information during training by equipping the critic with access to full plasma state information without diagnostics noise that is unavailable to the actor at inference time.<br>Tweaks and employed techniques include curriculum learning and plasma configuration randomization for better domain transfer.<br>To integrate this into operational settings, we implement real-time plasma control system (PCS) modules in C and C++ (depending on the target fusion device), ensuring performance, reliability, and compatibility with existing control infrastructure. |
| Testbeds or Pilot Deployments | More information on the pilot deployment can be found at [4] [5] |
| Code repositories | Not available |

# 2 Use Case Description

## 2.1 Description

Fusion energy is a promising clean and sustainable energy solution, offering significant potential for reducing greenhouse gas emissions on a large scale. Fusion Power Plants (FPPs) represent the next frontier in energy production. Flexible and robust control of FPPs requires managing the control object - plasma - under varying technical contexts, including diagnostic and material degradation, failures, and evolving conditions. Sustaining and efficiently controlling a steady-state fusion reaction in an FPP remains one of the most critical challenges in the fusion energy industry. Existing plasma control systems are designed to satisfy specific requirements of experimental devices, maintaining testing plasma discharges with finely tuned parameters. FPP will be operating in long discharges, guided by a reduced set of available measurements, which makes the conventional control approach less reliable and precise, creating a significant barrier to the commercial adoption of fusion energy.

Machine learning (ML) methods, particularly those based on reinforcement learning (RL), offer a promising alternative. To address this challenge, we are developing an innovative machine learning-based Plasma Control System (PCS) for both current and next-generation fusion devices, as well as future fusion power plants (FPPs). Our PCS is based on a machine-agnostic, plasma-state-oriented control approach and uses our in-house technology stack: an advanced NSFsim tokamak simulator and RL toolkit.

In this use case, we are developing the RL agent, a key component of PCS, for controlling plasma shape and position. This approach offers a promising alternative to conventional control methods by training end-to-end controllers that do not require additional data processing or task decomposition. The RL-based approach is faster to train and more adaptable to different plasma shapes, requiring no adjustments to training parameters, compared to conventional systems with PID controllers. During 2023-24, as part of the DIII-D National Fusion Facility Program, we applied RL algorithms for plasma parameter control on the DIII-D tokamak. Our

initial results demonstrate that privileged information enables the efficient training of ML agents while maintaining the simplicity of the Critic network's architecture. The Actor learns the control policy directly from sensor and actuator data, which can speed up the controllers by eliminating the intermediate step of magnetic flux distribution reconstruction. These advancements have laid the groundwork for developing essential components of the PCS. We believe that the intersection of ML and fusion science holds the key to unlocking sustainable, clean energy for our planet's future.

### Partners

- University of California, San Diego, USA
- General Atomics (DIII-D National Fusion Facility), USA

## 2.2 Benefits of the use case

1. In this use case, we use AI-based reinforcement learning and plasma state exploration to optimize plasma control, reduce disruptions, and enhance reactor efficiency. It helps increase the stability and energy output of fusion reactors, making fusion a more reliable and cost-effective clean energy source.
2. Our use case leverages advanced RL control modules and digital twin technology to modernize fusion reactor infrastructure. By enabling real-time plasma management and predictive diagnostics, our solution accelerates fusion commercialization and supports the development of next-generation energy.
3. We actively collaborate with leading research institutions, including the University of California in San Diego, Columbia University, University of Seville, and the DIII-D National Fusion Facility, as well as private fusion developers. By working with early adopters and industry stakeholders, we accelerate the validation, deployment, and scalability of AI-driven fusion control systems, contributing to the global clean energy transition.

## 2.3 Future Work

With additional resources, we plan to expand the list of controlled parameters and validate our ML-based Plasma Control System across a broader range of fusion devices in collaboration with early adopters. This would allow us to refine our models, enhance adaptability across different reactor designs, and further demonstrate the system's scalability and effectiveness. To achieve this, we would require more experiments and extensive data analysis, alongside the development of a simulation environment, which would enable us to train our ML models for more diverse control tasks and shift the control paradigm to control high-level plasma parameters, thereby defining overall plasma energy gain. This would improve the accuracy and robustness of real-time plasma state predictions and control strategies.

# 3 Use Case Requirements

**REQ-01**: Access to Experimental Data

It is critical that there is availability of time-series data from operating fusion devices (e.g., magnetic sensors, actuator signals) for training and validating ML control models.

**REQ-02**: Tokamak Simulator

It is critical that there is a physics-based, high-fidelity simulator (such as our NSFsim) to model plasma behaviour and evaluate control strategies in a tokamak environment.

TBD

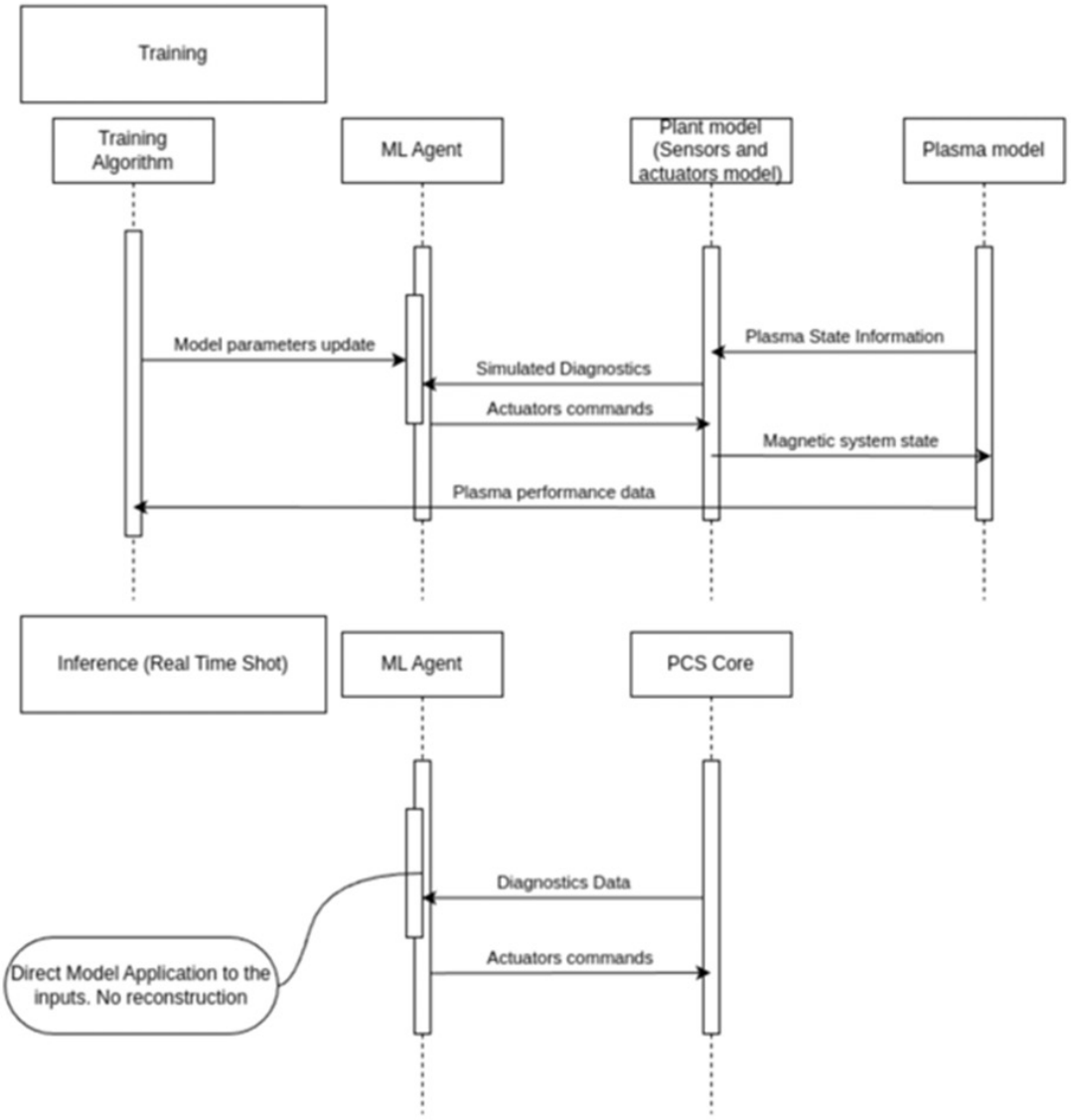
**REQ-03**: Machine-Agnostic Control Framework

It is critical that there is a flexible control architecture that can generalize across different fusion devices, plasma shapes, and operational scenarios without requiring manual reconfiguration.

**REQ-04**: Experimental Access

It is critical that there are opportunities to integrate and test the ML-based Plasma Control System in a real experimental setup to validate performance and robustness under real-world conditions.

## 4 Sequence Diagram



## 5 References

[1] Next Step Fusion, "NSFsim," [Online]. Available: https://blog.nextfusion.org/nsfsim-c5642b87aa47 (accessed: Jun. 15, 2025)

[2] DIII-D National Fusion Facility, "Become a User at the DIII-D National Fusion Facility," [Online]. Available: https://d3dfusion.org/become-a-user/ (accessed: Jun. 15, 2025)

[3] Next Step Fusion, "Building a Deep RL Model for Plasma Shape Control in Tokamaks," [Online]. Available: https://blog.nextfusion.org/building-a-deep-rl-model-for-plasma-shape-control-in-tokamaks-da2fa20a0775 (accessed: Jun. 15, 2025)

[4] G.F. Subbotin, D. et al. "Reconstruction-free magnetic control of DIII-D plasma with deep reinforcement learning." Available: https://arxiv.org/abs/2506.13267

[5] R. Clark et al. "Validation of NSFsim as a Grad-Shafranov Equilibrium Solver at DIII-D" Available: https://arxiv.org/html/2412.03786v1

# 4.6. Category 6: Finance

## Use Case 1: MOF IntelliServe – AI-Powered Conversational & Document Intelligence for Financial Services

UNITED ARAB EMIRATES
MINISTRY OF FINANCE
الإمارات العربية المتحدة
وزارة المالية

**Organization**: Ministry Of Finance

**Country**: United Arab Emirates

**Contact Person**:

Eng. Meshal Abdulla BinHussain, mabinhussain@mof.gov.ae

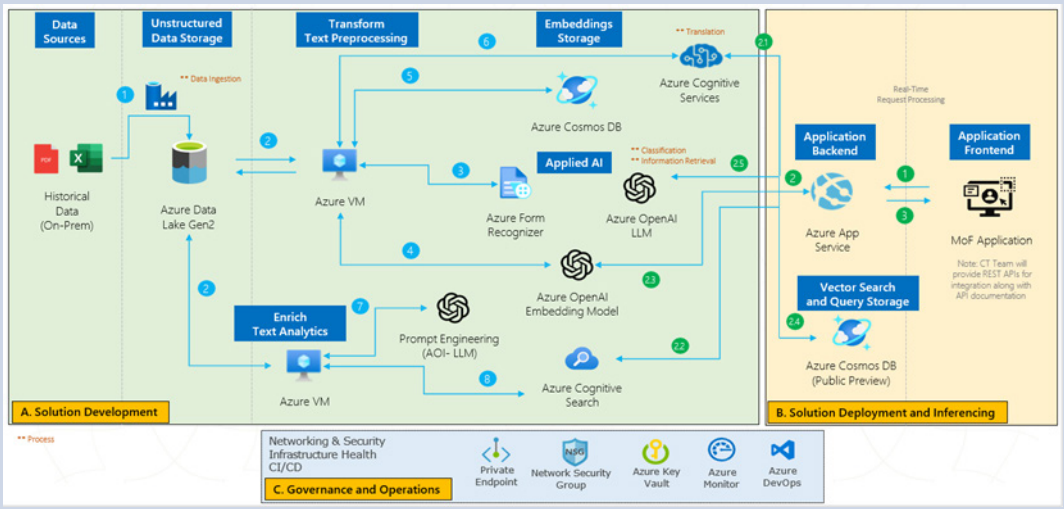
### 1 Use Case Summary Table

| Item | Details |
|---|---|
| Domain | Public Sector Financial Services & AI-driven Digital Transformation |
| Problem Addressed | • High Volume of Customer Calls<br>Customers call frequently to ask questions, which consumes significant call center time, as they struggle to find answers across various sources.<br>• Difficulty in Finding Information in Documents<br>Despite high downloads of documents, customers still complain about being unable to find the information they need within them. |

(continued)

| Item | Details |
|---|---|
| Key Aspects of Solution | • AI-Driven Conversational Chatbot<br>• An intelligent virtual assistant that responds to user queries related to MOF services, policies, and procedures in both Arabic and English.<br>• Uses advanced contextual memory, ensuring that follow-up questions are correctly interpreted and linked to previous interactions.<br>• Offers personalized real-time responses based on the user's intent, financial inquiries, and service requirements.<br>• Allows users to raise official requests if the chatbot response does not fully address their concerns.<br>• Persistent chat memory enables returning users to continue past conversations or start fresh based on their preference.<br>• AI-Powered Document Intelligence (Chat with PDF Feature)<br>• Enables users to upload financial documents (PDFs) and receive instant AI-generated summaries in both Arabic and English.<br>• Uses classifiers to determine if the document is relevant to MOF policies, laws, and financial regulations.<br>• Supports semantic search and real-time query resolution, allowing users to extract key insights from complex financial documents instantly.<br>• Employs Optical Character Recognition (OCR) to process scanned documents, ensuring accessibility to both digital and scanned files.<br>• Uses AI-**driven knowledge graphs** and **vector embeddings** to improve financial data extraction and response accuracy. |
| Technology Keywords | The **MOF IntelliServe** platform leverages cutting-edge AI technologies to enhance customer interactions and document processing. Key technologies include:<br>• Azure OpenAI Service<br>• AI-Powered Chatbot<br>• Natural Language Processing (NLP)<br>• Optical Character Recognition (OCR)<br>• Semantic Search<br>• Text Analytics<br>• Azure Cognitive Services<br>• Vector Search Algorithms<br>• Embedding Models<br>• Data Preprocessing<br>• Document Chunking<br>• Prompt Engineering<br>• Azure Cosmos DB<br>• Azure VM (Virtual Machines)<br>• AI-Driven Document Intelligence<br>• Real-Time Query Resolution<br><br>To conclude, the **MOF IntelliServe** platform is powered by a robust architecture, as shown in the **technical/functional diagram**, which outlines the seamless integration of Azure services, AI technologies, and data processing components that drive its performance and efficiency. |

(continued)

| Item | Details |
|---|---|
| |  |
| Data Availability | The **MOF IntelliServe** platform uses both **public** and **private** data sources to ensure comprehensive access to financial information.<br>• Public Data:<br>Data from publicly available sources such as government portals, public financial reports, and open data repositories.<br>• Private Data:<br>Internal financial data from the Ministry of Finance's secure databases, including financial records, and **service request history**.<br>These data sources are securely processed and integrated to ensure accurate, real-time responses for users. The platform ensures compliance with data security and privacy regulations. |
| Metadata (Type of Data) | The **MOF IntelliServe** platform handles various types of data to ensure accurate and efficient service delivery:<br>• Text Data:<br>Includes financial records, policies, user queries, and chatbot interactions.<br>• Visual Data:<br>Scanned documents (PDFs, images), including forms, policies, and financial statements processed using OCR.<br>• Metadata:<br>Includes document properties like title, author, and creation date, along with interaction logs such as query timestamps and user feedback. |

(continued)

| Item | Details |
|---|---|
| Model Training and Fine-Tuning | The **MOF IntelliServe** platform leverages the following steps for model training and fine-tuning:<br>**Data Collection**:<br>MOF-specific data is collected and prepared to train OpenAI models for better contextual understanding.<br>**Preprocessing**:<br>Data is cleaned and structured to ensure compatibility with OpenAI's pre-trained models, optimizing them for further customization.<br>**OpenAI Model Selection:**<br>OpenAI's pre-trained models are selected for their ability to handle complex tasks, making them ideal for this application.<br>**Fine-Tuning**:<br>The models are fine-tuned using MOF data and hyperparameters to improve accuracy and relevance in responses.<br>**Evaluation & Testing**:<br>The fine-tuned models are validated using performance metrics and accuracy checks to ensure optimal results.<br>This approach ensures that the AI models are customized to meet MOF's specific needs, offering accurate and efficient performance. |
| Testbeds or Pilot Deployments | The platform is currently in full production and is actively available for the Submit inquiries about the Ministry's services. It is fully deployed and operational, providing AI-driven services to users for seamless interaction and document processing. |

# 2 Use Case Description

## 2.1 Description

Customers of the Ministry of Finance (MOF) have historically faced challenges in efficiently accessing financial information. With multiple service portals, lengthy documents, and varying data formats, users often struggle to find the specific information they need quickly. This has led to an increase in calls to the call center, which is overloaded with repetitive inquiries, further hindering operational efficiency.

Moreover, many customers have difficulty extracting relevant details from large, complex financial documents. Despite downloading these documents, users are unable to easily locate the specific content they are looking for, creating frustration and delays in obtaining critical financial information.

These inefficiencies not only impact customer satisfaction but also increase operational costs for the Ministry, as employees spend significant time addressing repetitive queries and manually navigating through documents.

The primary objective of the MOF IntelliServe platform is to improve user access to financial information and services while enhancing operational efficiency. The key aims of this use case include:

1. Streamlining Customer Interaction:

   Provide users with an intuitive, AI-driven chatbot that delivers instant, accurate responses to their inquiries, eliminating the need for multiple service portals and long document searches.

2. Reducing Call Center Dependency:

   Automate responses to repetitive inquiries and route more complex requests to human operators, easing the workload of the call center and enabling faster resolution for customers.

3. Enhancing Document Accessibility:

   Enable users to upload and query financial documents directly, using AI to summarize and extract key information, making it easier to retrieve relevant details quickly and accurately.

4. Improving Overall Efficiency

   By automating document processing and customer interactions, the platform aims to reduce manual effort, streamline workflows, and cut down on operational costs for the Ministry of Finance

- Describe the innovative technological approach being implemented.

The **MOF IntelliServe** platform utilizes cutting-edge AI and machine learning technologies to address the challenges of customer interaction and document retrieval. The key technological innovations include:

1. **AI-Powered Chatbot**:

   An intelligent virtual assistant integrated with Natural Language Processing (NLP), allowing users to interact in both Arabic and English. The chatbot understands and responds to user queries in real-time, providing context-aware answers and guiding users to the right information efficiently.

2. **Chat-PDF Feature**:

   This unique feature allows users to upload financial documents (such as PDFs) and interact with them using the AI chatbot. The platform uses Optical Character Recognition (OCR) and semantic search to instantly extract and summarize relevant information from these documents, making it easier for users to find specific details.

3. **Data Processing and Embedding Models**:

   The system leverages advanced embedding models to process and classify data, enabling the AI to understand and answer complex queries. Azure OpenAI is used to power the backend, ensuring that the platform can scale efficiently while providing accurate and fast results.

4. **Real-Time Query Resolution**:

   By combining AI-driven document intelligence with vector search algorithms, users can query documents instantly, retrieving the most relevant answers in real-time. This reduces the need for manual document searches and allows for immediate access to crucial financial data.

These innovations enable **MOF IntelliServe** to not only automate responses but also enhance document management, ultimately improving the user experience and operational efficiency.

The MOF IntelliServe platform is designed to significantly enhance the Ministry of Finance's operations and customer service. The key impacts and benefits include:

1. **Improved User Experience**:

   By providing instant, AI-powered responses through the chatbot and enabling real-time document querying, users can quickly access the information they need, reducing frustration and improving satisfaction.

2. **Increased Efficiency**:

   Automating routine customer interactions and document processing reduces manual effort, allowing staff to focus on more complex tasks and increasing overall operational efficiency.

3. **Cost Savings**:

   With reduced dependency on call centers and the automation of document handling, the Ministry can expect a reduction in operational costs, allowing resources to be allocated more effectively.

4. **Scalability and Flexibility**:

   The use of Azure OpenAI and cloud infrastructure allows the platform to scale seamlessly, adapting to increasing data volumes and user demands without sacrificing performance.

5. **Data-Driven Insights**:

   Continuous data analysis from AI interactions and user feedback helps refine the system, providing valuable insights for future improvements in services and enhancing decision-making.

6. **Enhanced Accessibility**:

   With support for both Arabic and English, the platform ensures that financial services are accessible to a wider demographic, promoting inclusivity and meeting the needs of diverse users.

## Partners

**Microsoft**:

Microsoft plays a crucial role in providing the infrastructure and technical support for the project. As a key technology partner, Microsoft ensures the seamless deployment and operation of the solution by reviewing and validating the architecture and solutions before implementation begins. Their expertise also extends to supporting advanced tasks throughout the project, ensuring the integration of cutting-edge technologies and best practices.

Key Roles:

- Providing Azure infrastructure and cloud services.
- Reviewing the solution architecture and pre-implementation plans.
- Offering technical support during advanced tasks and troubleshooting.
- Ensuring smooth integration of OpenAI technologies into the solution.

## Celebal Technologies

Celebal Technologies is responsible for the end-to-end implementation of the AI-driven solution. They bring deep expertise in AI and machine learning to deliver a fully functional solution that automates customer inquiries and document processing. Celebal Technologies ensures the solution meets all client requirements, from development to deployment, and offers continued support post-implementation.

Key Roles:

- Designing and developing the AI-powered chatbot and document intelligence system.
- Implementing solution components, including integrations and automation.
- Conducting system testing and fine-tuning to optimize performance.
- Providing post-implementation support and continuous improvement services.

## 2.2 Benefits of the use case

- AI-driven automation enhances workforce productivity by reducing manual service inquiries and document processing, allowing government employees to focus on high-value tasks.
- MOF IntelliServe exemplifies AI-driven public service transformation, demonstrating how innovation in AI and automation enhances government financial services.
- Multilingual support (Arabic & English) ensures inclusive access to financial services, removing language barriers for different demographics.
- The integration of AI-driven automation into public financial services supports smart city initiatives, contributing to efficient digital governance.
- Enhances transparency and accountability in financial services by providing AI-powered document analysis and instant regulatory insights.
- Encourages collaboration between MOF, AI research institutions, and technology providers to enhance AI-driven financial services and governance.

## 2.3 Future Work

Looking ahead, several enhancements and developments are planned for the MOF IntelliServe platform:

1. **Expanded AI Model Integration**:

   Explore and compare other AI models beyond OpenAI, evaluating alternatives that could offer unique capabilities or better performance for specific use cases.

2. **Additional Features for Enhanced User Experience**:

   Future work will focus on introducing new functionalities, such as real-time language translation for more languages, sentiment analysis, and personalized recommendations based on user behavior and query patterns.

3. **Advanced Document Processing**:

   Further enhancements to the Chat-PDF feature will be implemented to improve document summarization and classification, offering even more granular insights into financial data.

- Outline additional resources needed, such as technology upgrades, further data collection, model improvements, etc.

To achieve these enhancements, the following resources will be required:

1. **Technology Upgrades**:

   We will continue to leverage Azure OpenAI but will also evaluate additional AI models to ensure the platform remains flexible and scalable. This may involve integrating new APIs, cloud resources, and data storage solutions.

2. **Further Data Collection**:

   Expanding the data sets with more diverse financial documents and user interaction data will be essential for further fine-tuning AI models and improving accuracy.

3. **Model Improvements**:

   Continuous improvement of AI models, particularly in fine-tuning them for specific user queries, will require dedicated resources for model training, performance testing, and validation.

Looking towards future collaborations and expansions:

1. **Academic and Research Collaborations**:

   We will explore partnerships with AI research institutions to stay at the forefront of emerging AI technologies and methodologies. This could involve research collaborations on natural language understanding, document processing, and AI ethics.

2. **International Expansion**:

   The system could eventually be adapted for use by other government entities or in different regions, leveraging the technology to meet broader public service needs.

## 3 Use Case Requirements

**Technical Requirements:**

- **REQ-01**: AI Model Integration

  Ability to integrate OpenAI and other AI models for comparison and flexibility in providing optimal solutions.

- **REQ-02**: Cloud Infrastructure

  Leverage Azure OpenAI services and other cloud resources for scalability, storage, and processing of large datasets and AI tasks.

- **REQ-03**: Data Security & Compliance

  Ensure the system adheres to UAE's data protection regulations and other relevant international standards for securing user data and sensitive financial information.

- **REQ-04**: Real-Time Processing

  The platform must be capable of processing user queries and documents in real-time, ensuring that responses and document retrieval are immediate.

**Operational Requirements:**

- **REQ-05**: Multilingual Support

  The system must support both Arabic and English, allowing users to interact seamlessly in either language.

- **REQ-06**: Scalability & Performance

  The solution should be able to scale with increasing user numbers and handle large volumes of data without affecting performance.

- **REQ-07**: 24/7 Availability

  The platform must be available 24/7, with high uptime to ensure continuous user access to the AI chatbot and document services.

**Functional Requirements:**

- **REQ-08**: AI-Powered Chatbot

  The system must include an AI-driven chatbot capable of understanding user queries through Natural Language Processing (NLP) and providing context-aware responses.

- **REQ-09**: Chat-PDF Integration

  The system must allow users to upload financial PDFs and enable them to interact with the content using AI-powered summarization, OCR, and semantic search.

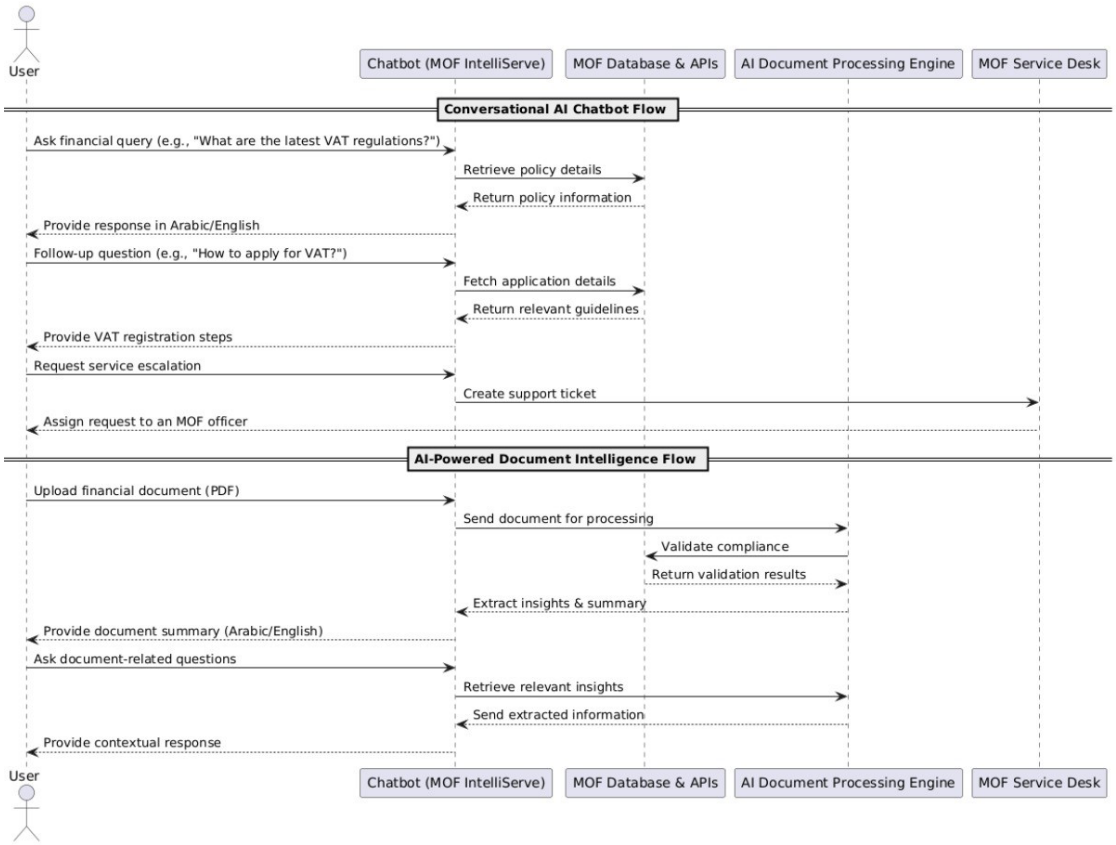
- **REQ-10**: Feedback Mechanism

  Implement a mechanism for users to provide feedback on AI responses, which will be used to improve the system through continuous learning.

- **REQ-11**: Real-Time Query Resolution

  The platform must resolve queries instantly, with accurate answers derived from structured and unstructured data sources.

## 4 Sequence Diagram



## 5 References

[1] Digital Dubai. "Dubai Digital Transformation Strategy." Accessed June 24, 2025. https://www.digitaldubai.ae.

[2] Ministry of Finance UAE. "Ministry of Finance UAE." Accessed June 24, 2025. https://www.mof.gov.ae.

[3] Telecommunications and Digital Government Regulatory Authority (TDRA). "UAE Data Protection Law." Accessed June 24, 2025. https://www.tdra.gov.ae.

[4] UAE Government. "UAE AI Strategy 2031." Accessed June 24, 2025. https://ai.gov.ae.

[5] World Economic Forum. "AI in Public Services Case Study." Accessed June 24, 2025. https://www.weforum.org.

## Use Case 2: AI-Powered Price Negotiation Solution (Nego-Wiz)

**Organization**: AI O2O

**Country**: Republic of Korea

**Contact Person(s)**:

Haewon Kam / haewonkam@o2o.kr

O2O Leaders Group / o2oteam@o2o.kr

## 1 Use Case Summary Table

| Item | Details |
|---|---|
| Category | Finances |
| Problem Addressed | Traditional B2B procurement negotiations are complex, time-consuming, and inconsistent, resulting in suboptimal business performance, high management costs, and limited use of data and real-time insights. |
| Key Aspects of Solution | AI-powered negotiation chatting system automates and standardizes repetitive negotiation tasks<br>• Real-time data analysis and benchmarking<br>• Adaptive negotiation personas for scenario-based strategies<br>• Integration with unified commerce platform |
| Technology Keywords | AI price negotiation, chat system, procurement automation, NLP, data analytics, persona modeling |
| Data Availability | Private (internal procurement and transaction data; no public repository) |
| Metadata (Type of Data) | Text (negotiation logs, product descriptions), numerical (pricing, transaction records), image (product images for benchmarking) |
| Model Training and Fine-Tuning | Utilizes large language models (LLMs) fine-tuned on procurement negotiation scenarios and enterprise transaction data. Adaptive persona logic for negotiation strategies. |
| Testbeds or Pilot Deployments | Pilot deployment within KT commerce's transaction platform (2024.12 ~2025.03)<br>Beta Launch Currently applied in real business operations for procurement negotiations(2025, 04)<br>Commercialization has been completed through a formal adoption agreement, initially covering a subset of product categories; expansion to encompass all transaction items of KT Commerce has been confirmed(2025.06). |
| Code repositories | Not publicly available(internal proprietary solution) |

KT Commerce is a subsidiary of KT Corporation, South Korea's largest and most prominent telecommunications company, providing advanced digital commerce and AI-powered procurement solutions that help businesses streamline online transactions and automate B2B negotiations.

# 2 Use Case Description

## 2.1 Description

AI Price Negotiation Solution, Nego-Wiz has been developed to revolutionize the process of corporate price negotiations [7]. Traditionally, e-commerce negotiations were time-consuming and inefficient, placing significant pressure on the MD (Merchandise Director) staff. The AI solution aims to address key challenges such as high transaction time, inefficiency in price negotiations, and the overburdening of MDs with manual tasks.

AI Price negotiation in this solution means using AI to automatically discuss and agree on the best price between buyers and sellers, making the process faster, fairer, and less work for people who used to do it manually.

### Problem Overview:

1. Transaction Time: Time spent on a single transaction was restricted by MD work schedules and labor-intensive negotiation processes. MDs would spend significant hours on negotiations, affecting their overall productivity.
2. Negotiation Efficiency: The manual process led to repetitive, inefficient negotiations, which were highly prone to errors due to human biases and emotional factors.
3. Negotiation Outcomes: MDs were unable to achieve optimal negotiation outcomes consistently, especially in complex or bulk transactions, as they were burdened by time and a lack of automated support.
4. Limited Data Utilization & Individually Skewed Insights: Negotiations often relied on fragmented or incomplete data, and information or insights were frequently biased toward individual perspectives, limiting organization-wide optimization and reducing fairness and transparency.

Solution Overview: The AI-powered solution automates price negotiation, reducing time and increasing efficiency while optimizing the negotiation outcome. By using AI, the solution:

1. Provides real-time AI-powered price negotiation that operates 24/7, automating repetitive tasks such as renegotiations and price adjustments based on real-time data.
2. AI personas in Nego Wiz adjust their negotiation tactics in real time by continuously ingesting and analyzing incoming negotiation data such as the partner's last offer, current market prices, historical deal outcomes, and defined target margins. Specifically, a "Self-Reflection" agent evaluates the short-term memory (STM) (i.e., the live negotiation context) against long-term memory (LTM) (i.e., past partner behavior and results) to select or modify a persona (e.g., cooperative, competitive, or neutral). That persona then drives the "Actor" agent to generate tailored responses (e.g., concession patterns or firm counteroffers). As new data (e.g., a sudden price cut by the supplier, changes in order volume, or shifts in partner responsiveness) enters STM, the Self-Reflection agent re-scores which persona and strategy combination is most likely to succeed, and the Actor agent updates its messaging accordingly [1][2]. This loop—data → persona adjustment → strategy execution → new data—allows the AI to dynamically shift its persona and tactics to maximize the probability of a favorable outcome. Integrated seamlessly with KT

Commerce's platform, enhancing negotiation capabilities by aligning real-time market prices with vendor responses [7].

(Short Version) AI personas in Nego-Wiz continually update by comparing live negotiation inputs (e.g., partner's last offer, market prices) stored in short-term memory against past outcomes in long-term memory. A "Self-Reflection" agent re-evaluates which persona (cooperative, competitive, neutral) best fits the evolving context. That persona then guides the "Actor" agent's responses, such as making a concession or holding firm. As each new offer or data point arrives, the system re-scores and switches personas as needed, ensuring tactics adapt dynamically to maximize success [4].

3. Aggregates and analyzes real-time data from diverse sources across the organization, delivering unbiased, organization-wide insights and recommendations that reduce individual bias, support data-driven negotiation strategies, and enable consistent, optimal decision-making for all stakeholders.

### Expected Benefits:

- The solution drastically reduces MD workload by automating price negotiation, allowing MDs to focus on more strategic tasks.
- Enhanced Negotiation Outcomes: Access to comprehensive, organization-wide insights allows for more strategic and effective negotiation strategies, leading to improved pricing, higher success rates, and greater overall business value.
- Increased Efficiency and Productivity: Automating data analysis and insight generation reduces manual effort, accelerates decision-making processes, and allows staff to focus on higher-value tasks.
- Greater Transparency and Collaboration: Organization-wide data integration fosters transparency and enables collaborative negotiation, ensuring all relevant perspectives are considered and supporting more equitable business outcomes.
- Unbiased, Organization-Wide Decision-Making: By aggregating and analyzing real-time data from diverse sources, the solution eliminates individual bias and delivers consistent, data-driven insights, enabling all stakeholders to make optimal decisions that benefit the entire organization.
- Continuous Improvement: The AI system learns from every negotiation, continuously refining its recommendations and strategies, which leads to ongoing optimization of negotiation processes and outcomes across the organization.

As a result, KT Commerce has successfully automated and optimized price negotiations, leading to a significant reduction in negotiation time, better alignment of product pricing, and negotiation success rate.

**Partners**

KT Commerce provides of 60,000+ annual transaction datasets and AI negotiation PoC verification.

Korea University HI-AI Laboratory provided advisory and support for the development of core technologies in artificial intelligence and natural language processing.

## 2.2 Benefits of the use case

- Industry, Innovation, and Infrastructure The solution uses AI-driven negotiation to modernize the way KT Commerce operates, promoting innovation in the e-commerce industry. The introduction of advanced AI technology creates a more efficient and

resilient business infrastructure, helping to drive growth and sustainability across the organization[6].

## 2.3 Future Work

### Planned Enhancements

1. Implementation of Multi-AI Agent System for Dynamic Strategy Adjustments: Future versions will enhance the Multi-AI Agent System, refining how the AI negotiates by leveraging multiple agents working in parallel to evaluate and execute various strategies in real time. This system will use a combination of specialized agents, each focusing on different aspects of the negotiation, such as price adjustments, competitor analysis, and partner response modeling. The Multi-AI Agent System will continuously adapt its approach based on the evolving negotiation dynamics, ensuring that each decision made aligns with both short-term goals and long-term strategy [5].

2. Advanced Goal Setting Mechanism: The Goal Setting function will be enhanced to allow the AI to establish dynamic, real-time target prices based on market conditions, competitor pricing, and transaction history. The AI will be able to automatically calculate the optimal target price, considering factors such as product value, negotiation history, and the negotiation goals set by the user. This will improve the precision of AI's pricing strategies and ensure that negotiation targets are always aligned with market realities.

3. Advanced Indicator System for Real-Time Strategy Assessment: The Indicator system will be refined to provide more granular real-time analysis of negotiation success probabilities. This will include visualizing not just the success rate but also dynamic factors like time, partner strength, and competitor pricing. The indicators will provide actionable insights, helping both the AI and human negotiators make informed decisions.

4. Integration of Real-Time Data Sources for Improved Price Prediction: The Lowest Price Search feature will be enhanced to incorporate real-time updates from a wider range of suppliers and marketplaces, ensuring that the AI can always propose the most competitive pricing during negotiations. This will also include expanding the use of big data and machine learning algorithms to predict price trends and adjust negotiations accordingly.

5. Expansion of Dynamic Negotiation Scenarios: As the AI learns from ongoing negotiations, it will expand its capability to handle more complex and diverse negotiation scenarios. Future updates will introduce additional negotiation variables like multiple negotiation rounds, bulk discounting strategies, and contract term negotiations. This will allow the AI to handle more varied and sophisticated negotiation processes.

### Resource Needs

1. Data Expansion and Integration: The AI will require access to an expanded dataset for improved learning and model training. This includes integrating data from additional market sources, transaction histories, and real-time competitor pricing. A stronger focus will be placed on obtaining high-quality transactional data to further train the AI for better decision-making during price negotiations.

2. Computational Power for Scaling: As the system becomes more advanced, the need for enhanced computational resources will grow. More powerful cloud infrastructure and data processing capabilities will be necessary to manage the increased volume of real-time data, complex multi-agent systems, and advanced machine learning algorithms.

3. Human Expertise in Continuous Improvement: Human input will continue to be valuable for tuning the AI's negotiation strategies and ensuring its adaptability. Experts in negotiation and AI model development will need to review the outcomes and fine-tune the system periodically. Additionally, feedback loops from real-world negotiators will ensure that the AI's performance aligns with industry standards.

Potential Future Collaborations

1. Industry Partnerships for Data Sharing and Expansion: Hyundai Motors, Byte Dance, KT(Korea Telecom)

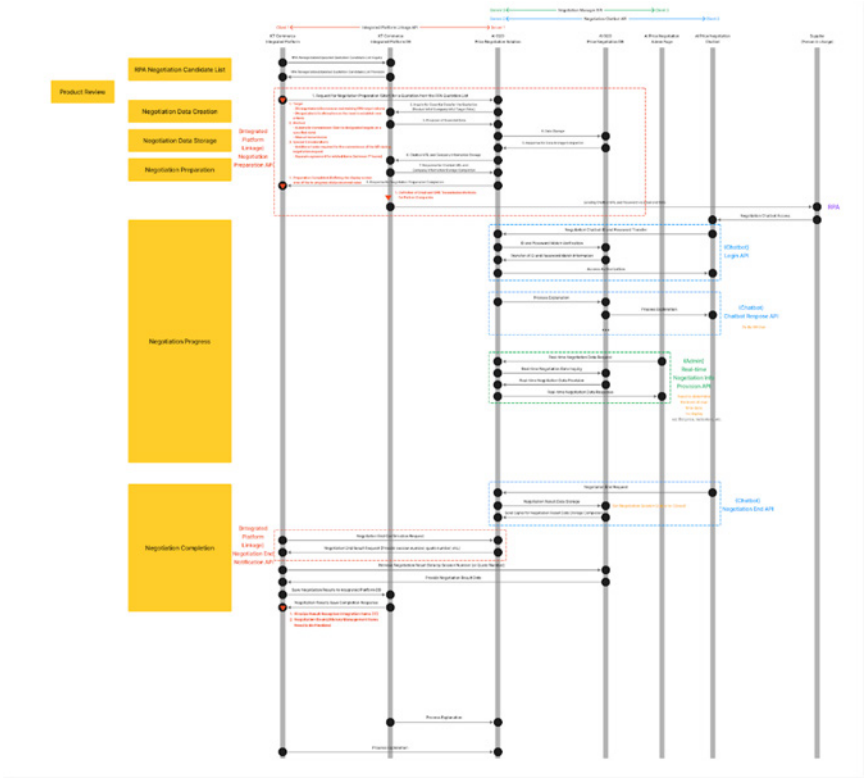
2. Collaborations with Academic and Research Institutions: Ongoing research collaborations, particularly with Korea University HI-AI Research and other institutions, will drive the development of advanced natural language processing (NLP) technologies and machine learning models. These collaborations will also introduce the integration of Long-Term Memory (LTM) and Short-Term Memory (STM) in the negotiation system [1]. By applying LTM and STM, the AI will have enhanced capabilities to retain and recall negotiation history, which will improve its decision-making processes. LTM will store long-term strategic learnings and insights gained from past negotiations, while STM will focus on immediate negotiation data, allowing the AI to respond more effectively to ongoing interactions. This memory-based approach will enable the AI to adapt strategies based on both immediate and historical contexts [4].

## 3 Use Case Requirements

Technical Requirements:

- **REQ-01**: It is critical that the system be implemented as well as deployed to seamlessly integrate with KT Commerce's existing back-end database to facilitate real-time data exchange and ensure smooth operation across platforms.
- **REQ-02**: It is expected that the system be implemented but may not be deployed to support real-time communication between the AI and human negotiators, ensuring that the negotiation progresses without delays and adapts to changes in real time.
- **REQ-03**: It is critical that the system be implemented as well as deployed to offer an intuitive and easy-to-navigate interface for both MDs and customers to engage in price negotiations efficiently.
- **REQ-04**: It is of added value that the AI be capable of dynamically adjusting its persona based on the negotiation data, such as the negotiation partner's profile, market conditions, and past negotiation outcomes. This need not be implemented or deployed immediately.

## 4 Sequence Diagram




## 5 References

[1] S. Yao et al., "ReAct: Synergizing Reasoning and Acting in Language Models," National Science Foundation (NSF) Public Access Repository. [Online]. Available: https://par.nsf.gov/biblio/10451467

[2] J. Liu et al., "Reflexion: Language Agents with Verbal Reinforcement Learning," in Advances in Neural Information Processing Systems (NeurIPS), 2023. [Online]. Available: https://proceedings.neurips.cc/paper_files/paper/2023/hash/1b44b878bb782e6954cd888628510e90-Abstract-Conference.html

[3] Z. Zhu et al., "NegotiationArena: How Well Can LLMs Negotiate? Platform and Analysis," arXiv preprint, arXiv:2402.05863, 2024. [Online]. Available: https://arxiv.org/abs/2402.05863

[4] H. Shen et al., "Improving Language Model Negotiation with Self-Play and In-Context Learning from AI Feedback," arXiv preprint, arXiv:2305.10142, 2023. [Online]. Available: https://arxiv.org/abs/2305.10142

[5] A. Author(s), "Proceedings of the AAAI Conference on Artificial Intelligence," AAAI, vol. 38, no. 5, 2024. [Online]. Available: https://ojs.aaai.org/index.php/AAAI/article/view/29936

[6] FasterCapital, "10 Ways AI for Inventory Optimization is Reducing Costs," [Online]. Available: https://fastercapital.com/articles/10-Ways-AI-for-Inventory-Optimization-is-Reducing-Costs.html

[7] NEGO WIZ, "Introducing NEGO WIZ, AI Powered Price Negotiation Solution," YouTube, May 31, 2025. [Online]. Available: https://youtu.be/XRnlMUycY7s

# Use Case 3: AIZEN – CreditConnect-EV

[CreditConnect-EV]

**Organization**: AIZEN

**Country**: South Korea

**Contact Person(s)**:

Jihoon Lim, Jihoon.lim@aizen.co

Jung Seok Kang, js.kang@aizen.co

## 1 Use Case Summary Table

| Item | Details |
|---|---|
| Category | Finance |
| Problem Addressed | A persistent financing gap is slowing the conversion to electric twowheelers (E-2Ws) across Southeast Asia. Conventional lenders lack the granular, real-time data and analytic tools needed to price risk accurately, leaving many riders and fleet operators without affordable credit. |
| Key Aspects of Solution | Internet of Things(IoT)-driven data capture: Continuous telemetry from E-2W batteries (System on a Chip(SoC), temperature, charge cycles) and on-bike sensors (speed, mileage, acceleration).<br>Behavioural analytics: Riding patterns, payment history, and geospatial usage profiles enrich credit signals.<br>Automated Machine Learning(AutoML) risk engine: Our proprietary AutoML pipeline ingests these high-frequency data streams, predicting residual value, default likelihood, and optimal loan terms in near real time.<br>Embedded financing: Insights flow directly into the CreditConnect-Electric Vehicle(EV) platform, enabling partner lenders to approve or price loans that would be infeasible under traditional scorecard models. |
| Technology Keywords | Big-data telemetry, AutoML modelling, Predictive risk scoring, Edge/Cloud IoT, Application Programming Interface(API)-first embedded finance |
| Data Availability | Private |
| Metadata (Type of Data) | Text |

(continued)

| Item | Details |
|---|---|
| Model Training and Fine-Tuning | Gradient-boosted trees and deep tabular networks trained via our AutoML stack<br>Self-supervised temporal transformers (e.g., Time2Vecenhanced Thin-Film Transistor(TFT)) pre-train on unlabeled battery/Rider IoT streams, then fine-tune for PD (default probability) and RV (residual value) with <1 % label use.<br>Heterogeneous graph neural networks (HGNNs) model rider–bike–route relations, capturing cross-entity contagion effects and localized battery-health clusters.<br>Category-adaptive federated learning syncs Original Equipment Manufacturer(OEM)-specific heads while sharing a common backbone, reducing cold-start error by ≈20 % when a new bike model is introduced.<br>Continual learning with weekly back-testing |
| Testbeds or Pilot Deployments | CreditConnect-EV live since 2022 in Indonesia: Over 20,000 E-2Ws financed to date (about one-third of all newly adopted E-2Ws) |

# 2 Use Case Description

## 2.1 Description

Southeast Asia's rapid-growth mobility sector relies heavily on gasoline-powered two-wheelers. In Indonesia alone, more than 120 million motorcycles are on the road, emitting over 58 million tons of $CO_2$ annually. Although electric two-wheelers (E-2Ws) can reduce operating costs by up to 40 % and eliminate tail-pipe emissions, adoption is constrained by the high upfront price and the lack of credit products tailored to this new asset class.

Traditional lenders depend on static income statements and collateral checks, neither of which capture the real-time performance and utilization profile of an E-2W. As a result, riders, fleet operators, and small delivery businesses—exactly the groups that would benefit most—remain locked out of affordable financing.

Objectives include:

- Bridge the financing gap by turning high-frequency battery and rider telemetry into alternative credit signals that enable banks and non-bank lenders to underwrite E-2W loans confidently.
- Accelerate EV conversion across ride-hailing, logistics, and e-commerce delivery segments, lowering CO2 and particulate emissions in congested urban corridors.
- Improve livelihoods for riders and small-business owners by reducing total cost of ownership and unlocking new income opportunities.
- Derisk lender portfolios through dynamic residual-value forecasting and continuous risk monitoring, lowering default rates and enabling sustainable interest pricing.

Use Case Status**:** This use case is currently in an active pilot stage across Indonesia, with real-world deployment of telemetry-enabled E-2Ws and risk-scoring engines integrated into lending workflows.

**Partners**: Grab, Gojek, Lazada, KB Capital, Woori Finance, Sinarmas Hana Capital, HSBC, Accial Finance, E-2W OEMs & Battery-swap Operators

## 2.2 Benefits of the use case

The project delivers significant benefits by financing electric two-wheelers (E-2Ws), which effectively displaces petrol bikes. This shift not only raises the share of clean energy in the transport sector but also cuts per-kilometer energy use, contributing to a more sustainable and energy-efficient transportation system. By deploying locally built IoT and AutoML technology, the project modernizes financial and mobility infrastructure, ensuring that these systems are upgraded for sustainability. Additionally, the project enables low-emission delivery and ride-hailing fleets, which helps to lower particulate emissions in dense urban areas. This not only improves air quality but also enhances the overall sustainability of urban transportation, making cities healthier and more livable.

## 2.3 Future Work

To further enhance our initiative and expand its impact, we plan to extend our pilots to Vietnam and the Philippines, adapting our models to align with local riding patterns and regulatory contexts. This geographic expansion will ensure that our solutions are tailored to meet the specific needs of these regions, maximizing their effectiveness and relevance.

We will also integrate battery-swap station data, including cycle counts and dwell time, into our systems to refine residual-value estimates. This integration will provide more accurate and reliable data, enhancing our ability to manage and optimize battery usage and performance.

To improve efficiency and resilience, we will deploy lightweight AI models on IoT gateways, enabling sub-100 ms scoring and ensuring offline capabilities. This edge AI scoring will significantly enhance our operational speed and reliability, even in areas with limited connectivity.

We are introducing Environmental, Social, and Governance(ESG)-linked loan products, offering rate discounts tied to emission-reduction milestones verified by telemetry. This innovative approach not only promotes environmental sustainability but also incentivizes users to adopt greener practices, contributing to a more sustainable future.

To enhance our feature set, we will partner with telecommunications companies and ride-hailing firms to incorporate telecom spend and trip-density features. These additional data points will improve the predictive power of our models, providing more accurate and actionable insights.

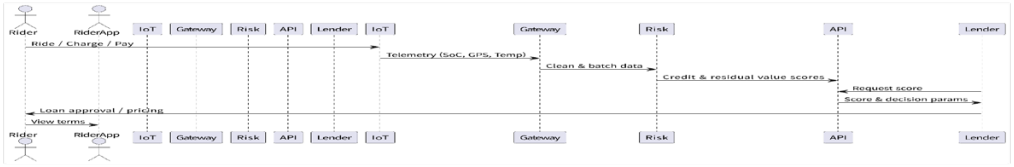
We will also upgrade our models by incorporating self-supervised contrastive pre-training on unlabeled telemetry data. This approach will boost accuracy even with sparse ground truth data, enhancing our ability to make precise predictions and decisions.

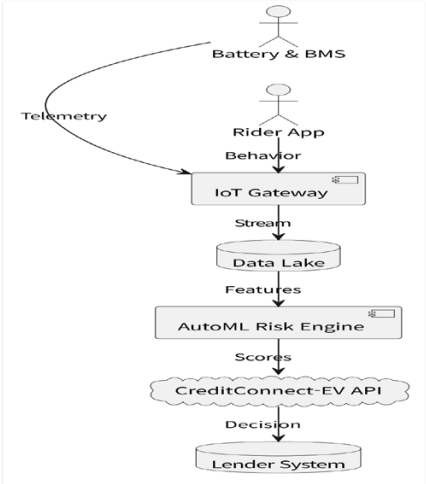
To support these advancements, we will need additional resources, including more Graphics Processing Unit(GPU) nodes for weekly retraining, expanded Long-Term Evolution(LTE)/5G bandwidth, and the addition of two senior data scientists to our team. These resources will ensure that we can continue to innovate and improve our services, delivering greater value to our users and stakeholders.

## 3 Use Case Requirements

- **REQ-01**: It is critical that the data pipeline is capable of ingesting at least 5 GB of data per day while maintaining an end-to-dashboard latency of less than 5 minutes. This ensures real-time visibility and responsiveness in data-driven decision-making.
- **REQ-02**: It is expected that the risk engine API delivers risk scores with a response time under 500 milliseconds at the 95th percentile (P95). This level of performance is necessary to support time-sensitive applications and user experiences.
- **REQ-03**: It is expected that machine learning models are retrained on a weekly basis to maintain relevance and accuracy. Additionally, it is critical that immediate retraining is triggered if data drift is detected, as indicated by a Kolmogorov–Smirnov (KS) statistic greater than 0.2.
- **REQ-04**: It is critical that all API endpoints comply with OpenAPI 3.0 standards and are secured using OAuth 2.0 protocols to ensure data integrity, authentication, and secure access.
- **REQ-05**: It is critical that IoT communication links are secure, encrypted, and achieve a minimum uptime of 99%, thereby ensuring reliability and resilience in continuous operations.

## 4 Sequence Diagram



Data Analysis using Google BigQuery Data is analyzed using BigQuery, a cloud-based data warehouse suitable for handling large datasets and complex Structured Query Language(SQL)-based queries.

Untitled query | Run | Save | Download | Share | Schedule | Syntax error: Unexpected identifi...

```
with cust as ( -- default asli 20 (inactive)
  select loan_id, status, motorcycle_id, date(activated_timestamp) activated_date, case when status = 20 then date(updated_at)
else current_date end default_date
  from `AIZEN_ID.LOAN_APPLICATION`
  WHERE GRASS_DATE = CURRENT_DATE - 1
  and deleted = 0
  and program_type = 0
  and status >= 16
  -- and date(updated_at) >= '2024-10-01'
)
, veh as (
  select _id, case when lower(vehicle_brand) = 'smoot' then vehicle_id else vehicle_vin end vehicle_id
  from AIZEN_ID.VEHICLE_PROFILE
  where grass_date = current_date - 1
  and program_type in (0,1,5)
)
, batt as (
```

Press Option+F1 for Accessibility Options.

Query results | Save results | Open in

Job information | Results | Chart | JSON | Execution details | Execution graph

| Row | vehicle_id | activated_date | default_date | total_active | total_swapping | is_default | s |
|---|---|---|---|---|---|---|---|
| 1 | MC-16052300210111 | 2024-04-18 | 2025-04-25 | 24 | 7 | 0 | 0. |
| 2 | MC-16052300206411 | 2025-04-17 | 2025-04-25 | 7 | 8 | 0 | 1. |
| 3 | MC-16052300206411 | 2025-02-25 | 2025-03-05 | 8 | 21 | 1 | |
| 4 | MC-16052300386511 | 2025-02-04 | 2025-04-25 | 78 | 155 | 0 | 1. |
| 5 | MC-16052300441411 | 2024-04-19 | 2025-04-25 | 191 | 215 | 0 | 1. |
| 6 | MC-16052300200011 | 2024-04-19 | 2025-04-25 | 190 | 186 | 0 | 0. |
| 7 | MC-16052300197111 | 2024-04-18 | 2024-11-06 | 11 | 47 | 1 | 4. |
| 8 | MC-16052300197111 | 2025-02-25 | 2025-04-25 | 58 | 71 | 0 | 1. |

Results per page: 50 | 1 – 50 of 92

Data Visualization using Metabase Metabase is used to create dashboards and visual representations of the data. This enables stakeholders to derive insights and make data-driven decisions effectively.

## 5 References

[1] Zhang, S. et al. State of health prediction for lithium-ion battery using a gradient boosting-based data-driven method. 2024

[2] World Bank, Recommendations to the National Roadmap and Action plan for the Electric Mobility Transition, 2023, OJK, Digital Finance

# Use Case 4: AI-Powered Payroll Solution (AI Payroll Guard)

UNITED ARAB EMIRATES
MINISTRY OF FINANCE

الإمارات العربية المتحدة
وزارة المالية

**Country**: United Arab Emirates

**Organization**: Ministry of Finance

**Contact Person**:

Shaikha Rashed Musabbeh Saeed Alkaabi, sralkaabi@mof.gov.ae

## 1 Use Case Summary Table

| Item | Details |
|---|---|
| Category | Finance |
| Problem Addressed | This AI-powered system automates repetitive tasks, reduces human errors, ensures accuracy in salary disbursements, and enhances transparency. |
| Key Aspects of Solution | This AI-powered system has significantly reduced payroll processing time by up to 95%, ensuring compliance with UAE legislation. These improvements also align with the UAE government's sustainability and innovation goals, as AI-driven automation reduces resource consumption and streamlines processes |
| Technology Keywords | Artificial Intelligence (AI), Robotic Process Automation (RPA) |
| Data Availability | Private |
| Metadata (Type of Data) | Employee Data (e.g., employee status, grade, bank info etc.), Transaction Data (e.g., Compensation, benefits values calculated), Historical payroll data to train the AI |
| Model Training and Fine-Tuning | Llama 3.1 (8B), UI Path (RPA), Oracle (Payroll), Impala (Data Lake), Python Flask (Chatbot) |
| Testbeds or Pilot Deployments | aipayrolldev.mof.gov.ae |

## 2 Use Case Description

### 2.1 Description

The Ministry of Finance (MoF) of the UAE has successfully implemented an innovative AI-enabled payroll system, marking a significant leap in the digital transformation of federal financial operations. This cutting-edge system leverages the power of Artificial Intelligence (AI) and Intelligent Automation (IA) to streamline and optimize payroll processing across all federal entities, delivering substantial improvements in efficiency, accuracy, and transparency.

The newly deployed system is designed to tackle the inherent complexities and time-consuming nature of traditional payroll management. By automating repetitive tasks and minimizing human intervention, it effectively reduces the potential for errors, thereby ensuring impeccable accuracy in salary disbursements. This robust approach to payroll management also significantly enhances transparency within the financial processes of the federal government.

Key functionalities of the AI-enabled payroll system include:

- Automated Monitoring and Processing: AI actively monitors payroll schedules, triggering IA agents to initiate and process payroll operations precisely when needed.
- Data Integrity and Validation: The system incorporates advanced AI capabilities to validate data integrity, ensuring that all payroll information is accurate and consistent.
- Exception Resolution: AI is instrumental in identifying and resolving exceptions within the payroll process, mitigating potential delays and discrepancies.
- Automated Salary Payments: The entire salary payment process is automated, from calculation to disbursement, ensuring timely and accurate payments.
- Chatbot-Enabled User Interface: An integrated chatbot provides a user-friendly interface for monitoring payroll progress and facilitating approvals, enhancing accessibility and control for relevant stakeholders.

The impact of this AI-powered transformation is profound. The MoF reports a remarkable reduction in payroll processing time by up to 95%. This drastic improvement not only boosts operational efficiency but also ensures strict compliance with all relevant UAE labor legislation and financial regulations.

Supervising an AI solution that processes payroll requires a balance of oversight, validations, and continuous improvement to ensure accuracy, compliance, and trust.

Status Monitor: - An admin console is used to monitor the status of payroll processes being executed by AI

Data Checkpoints: - AI-powered payroll systems automatically perform data validation by cross-referencing payroll data, detecting discrepancies, and flagging anomalies such as duplicate entries, missing data, or unusual salary amounts.

- When the AI system identifies exceptions or anomalies during payroll data validation, these are presented to payroll administrators or designated personnel for manual review and resolution before payroll is finalized.
- Manual review serves as a critical checkpoint to ensure accuracy, compliance, and data integrity.
- The process typically involves:
- Reviewing flagged exceptions in a dashboard or report generated by the AI system.
- Investigating the root cause of each exception (e.g., incorrect data entry, missing documentation, policy violations).

Exception Handling: - AI to notify administrators of uncertain technical errors for review

Human-in-the-Loop: - Critical approval as per the delegation of authority prior to salary disbursements

Alternative Plan: The Alternative plan can be initiated in case of any unforeseen events

Use Case Status: The system is live with Payroll Processing of Federal Ministries.

**Partners**

Government Stakeholders: Ministry of Finance (MoF), UAE Federal Entities.

## 2.2 Benefits of the use case

- Goal 9: Industry, Innovation, and Infrastructure – AI and IA are leveraged to develop an advanced payroll infrastructure.
- Goal 16: Peace, Justice, and Strong Institutions – The system enhances transparency, reduces fraud, and ensures compliance with financial governance frameworks.
- Goal 8: Decent Work and Economic Growth – By reducing manual work, the system enhances productivity and improves payroll accuracy.

## 2.3 Future Work

1. Payroll Inquiries Automation

Employees often have recurring payroll-related questions that require immediate answers. AI chatbots can handle such queries efficiently, including:

Salary Breakdown & Payslip Requests

- Employees can ask: "How much was my salary last month?" or "Can I get my latest pay slip?"
- The AI chatbot retrieves salary details from the payroll system and provides a downloadable pay slip.

Deduction Explanations

- Employees can inquire: "Why is my deduction higher this month?"
- The AI chatbot analyses reasons and explains the impact on net salary.

End of Service Benefits Query

- Example: "What is my gratuity?"
- AI can fetch analyses and fetch the gratuity data.

Payroll Processing Status

- If an employee asks, "Has my salary been processed?", the chatbot can provide real-time payroll status updates.

Banking & Payment Queries

- Example: "I changed my bank account. When will my salary be credited?"
- The bot checks updated bank details and payment timelines and provides an expected disbursement date.

Impact:

- Reduces HR workload on repetitive payroll queries.
- Ensures 24/7 availability of payroll support.

- Improves accuracy in responses, reducing payroll-related escalations.

2. Automated Employee Reimbursements

How It Works:

AI-powered systems streamline the employee reimbursement process by automatically verifying, approving, and processing expense claims related to travel, entertainment, and operational costs. Traditionally, reimbursements require manual review by finance teams, which can be time-consuming and error-prone. AI eliminates these inefficiencies by:

- Scanning Receipts with OCR (Optical Character Recognition) or the employees uploading receipts or invoices using a mobile app.
- AI extracts relevant details (amount, date, vendor, location) and matches them against policy rules.
- Validating Company Policies
- If an expense violates a policy, AI flags it and provides a reason for rejection.
- Auto-Approving Standard Claims (If a claim meets all policy criteria, AI approves it instantly without human intervention)
- Reduces processing time from days to minutes.
- Seamless Integration with Payroll & Finance Systems
- Approved claims are directly processed for reimbursement in the next payroll cycle.
- AI ensures correct tax deductions where applicable.
- Benefits: Faster Reimbursements Employees receive payments promptly, improving satisfaction.
- Reduced Workload for Finance Teams – No need for manual claim verification.
- Minimized Errors – AI ensures accurate data entry and compliance checks.
- Improved Transparency – Every claim is logged with an audit trail for accountability.

## 3 Use Case Requirements

Automated Payroll Release for Federal Ministries & Authorities

- **REQ-01**: End-to-End Payroll Automation

  It is critical that the system fully automates the payroll release process for Federal Ministries & Authorities, integrating AI and RPA technologies to manage the entire payroll lifecycle, from scheduling to execution, with real-time exception monitoring and resolution.

- **REQ-02**: Reduced Manual Intervention, Compliance, and Auditability

  It is expected that the automation will significantly reduce manual intervention, ensuring compliance and maintaining an audit trail, thereby enhancing the efficiency and accuracy of payroll operations (implemented but may not be deployed).

- **REQ-03**: Seamless System Integration

  The system shall integrate seamlessly with existing HR and financial systems to retrieve, process, ensuring data consistency, and eliminate duplication.

- **REQ-04**: AI-Driven Exception Handling

The system shall utilize AI algorithms to detect, monitor, and resolve payroll exceptions or anomalies in real time, with escalation to authorized personnel when automated resolution is not possible.

- **REQ-05**: Comprehensive Audit Trail

  The system shall maintain a comprehensive, immutable audit trail that records all payroll actions, changes, approvals, and exception resolutions for compliance and reporting purposes.

- **REQ-06**: Secure, Role-Based Access

  The system shall provide secure, role-based access controls, ensuring that only authorized personnel can view, approve, or modify payroll data and processes.

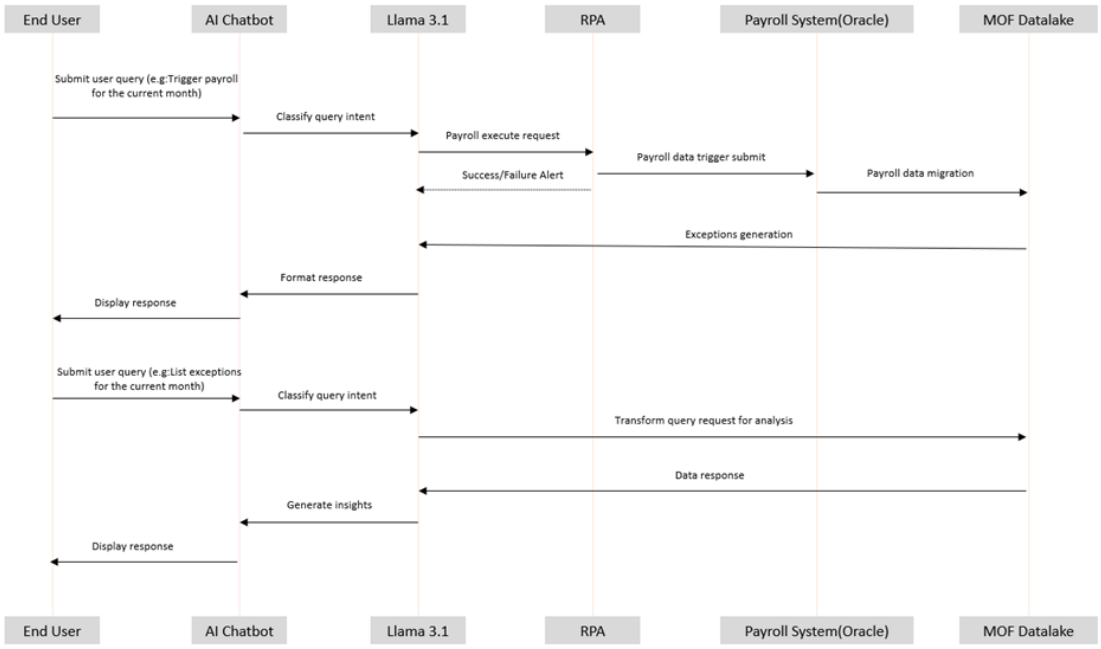
- **REQ-07**: Real-Time Dashboards and Reporting

  The system shall offer real-time dashboards and generate detailed reports on payroll status, exceptions, compliance metrics, and audit logs for transparency and oversight.

- **REQ-08**: Scalability and Adaptability

  The system shall be scalable to support additional ministries, authorities, or changes in payroll policies and regulations as required.

## 4 Sequence Diagram




## 5 References

[1] Ministry of Finance UAE. "*Video.*" Instagram, October 16, 2024. Accessed June 19, 2025. https://www.instagram.com/mofuae/reel/DBLmoYbyofv/.Ministry of Finance UAE.

[2] *AI Payroll Development Portal*. https://aipayrolldev.mof.gov.ae.

# Use Case 5: AI-Driven Financial Inclusion with Enhanced Consumer Protection via Mobile Networks in Zambia

**Organization**: Zambia Telecommunications Company (Zamtel)

**Country:** Zambia

**Contact Person**:

Paul Chiluba Cheche, paul.cheche@zamtel.co.zm, paulcche2@gmail.com

## 1 Use Case Summary Table

| Item | Details |
|---|---|
| Category | Finance |
| Problem Addressed | Limited access to formal financial services for Small and Medium-sized Enterprises (SMEs), farmers and marginalized populations; high vulnerability to mobile-money fraud and low financial literacy. |
| Key Aspects of Solution | AI-driven alternative credit scoring using anonymized mobile transaction data;<br><br>Real-time fraud detection & prevention via Machine Learning (ML) anomaly detection;<br><br>AI-powered Natural Language Processing (NLP) chatbots for financial literacy & dispute resolution. |
| Technology Keywords | AI, ML, NLP, Mobile Money, USSD, SMS, Data Analytics, Biometrics |
| Data Availability | Private |
| Metadata (Type of Data) | Zamtel Transaction logs, CDRs, app usage data, SMS/USSD content, demographic profiles, literacy assessment responses |
| Model Training & Fine-Tuning | ML Supervised classification/regression (credit scoring); anomaly detection (fraud); transformer-based NLP fine-tuned on English and Local language datasets |
| Testbeds or Pilot Deployments | Initial pilot within the Zamtel network |

# 2 Use Case Description

## 2.1 Description

In Zambia, 60% of adults remain unbanked due to lack of formal credit histories, geographic barriers and systemic biases (World Bank, 2020). Mobile money is widespread, yet fraud and illiteracy hinder its benefits.

The aim is to enable over 500,000 unbanked Zambians to access micro-credit; reduce SMS-based scam incidence by 60% and deliver financial education.

Innovative Approach include:

Credit Scoring: Explainable AI (logistic regression + Shapley Additive explanations (SHAP) on mobile transaction patterns (Talk time top-ups, utility payments, etc.). When financial distress is detected, the system triggers support actions like counselling and tailored credit suggestions to help the customer regain stability.

Fraud Detection: Real-time anomaly detection at network edge (Lower Latency, Reduced Backhaul Congestion, Enhanced Privacy etc.)

Chatbots: Localized NLP models (Masakhane) delivering Short Message Service/ Unstructured Supplementary Service Data (SMS/USSD) guidance in local languages; real-time dispute resolution via Zambia Information and Communications Technology Authority (ZICTA) hotline integration for enhanced compliance by service providers.

**Use Case Status:** The use case is part of a big project in progress

### Partners

Zamtel and Zambia Information and Communications Technology Authority(ZICTA)

## 2.2 Benefits of the use case

Our initiative delivers substantial benefits by expanding access to credit, which significantly boosts income-generating opportunities for individuals and communities. By providing microloans, we foster the growth of SMEs and create new job opportunities, driving economic growth and stability. Leveraging AI on existing mobile infrastructure, we enhance fintech innovation, making financial services more accessible and efficient. Our fair AI scoring system ensures that women and customers in rural areas, regardless of their economic status, have equal access to financial opportunities, thereby reducing inequalities. Proactive fraud detection mechanisms strengthen trust in institutions by ensuring the integrity and security of financial transactions. Furthermore, we foster collaboration with operators, regulators, and academia to create a robust ecosystem that supports sustainable development and inclusive growth.

## 2.3 Future Work

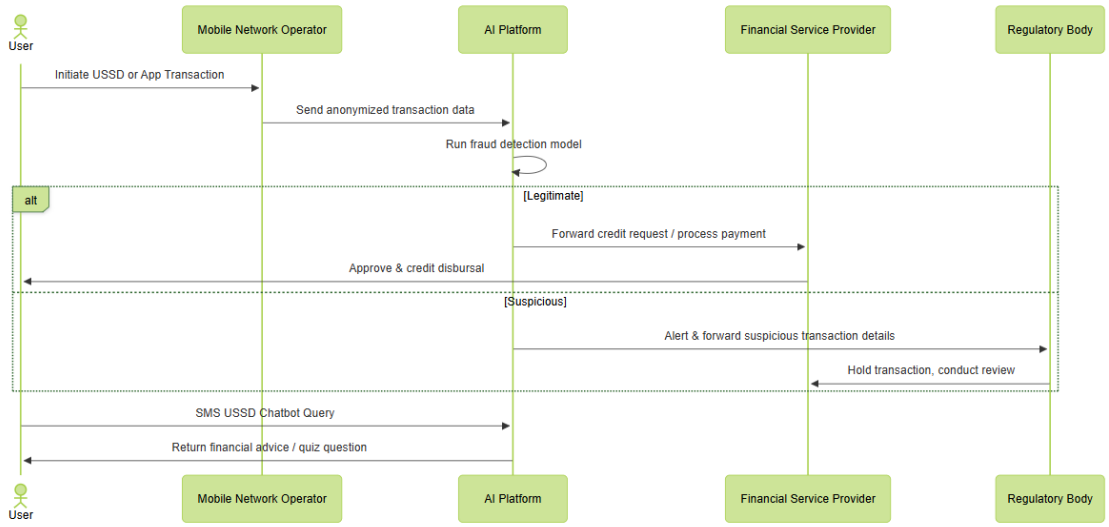
To further strengthen our platform, we plan to introduce several forward-thinking enhancements. These include biometric authentication for greater security and ease of use, AI-assisted financial guidance to help users make smarter financial decisions, and extended dialect support to make sure the platform is truly accessible and relevant across our region. To enable these

innovations, we will need more comprehensive datasets, especially from sectors like agriculture and utility payments, alongside advanced edge-AI hardware that can handle processing directly on low-cost mobile devices even in challenging network environments. We will also develop a companion mobile app to provide a more integrated, convenient, and user-centric experience. To bring this to life, we plan to collaborate with international development agencies for funding, while strengthening local and regional knowledge sharing with other mobile service providers and fintech innovators. This collaborative approach will enable us to leverage best practices and implement cutting-edge solutions tailored to the unique needs of our users. Lastly, we will partner with innovative FinTech startups to keep our platform at the forefront of financial technology ensuring we remain adaptable, forward-thinking, and directly aligned with the growing financial empowerment of our communities.

## 3 Use Case Requirements

- **REQ-01**: It is critical that secure and privacy-preserving data aggregation and anonymization protocols are implemented in partnership with telecom providers. These protocols must protect user identity while enabling reliable credit risk analysis.
- **REQ-02**: It is expected that robust AI and machine learning models for credit scoring and fraud detection are developed and fine-tuned to reflect local behavioral and transactional patterns in Zambia. This ensures that creditworthiness is assessed fairly and fraud risks are accurately detected.
- **REQ-03**: It is critical that the AI services are seamlessly integrated into existing Zamtel mobile money platforms, including USSD and SMS channels. This allows for broad accessibility, especially for users with limited digital literacy or smartphone access.
- **REQ-04**: It is of added value that localized natural language processing (NLP) modules are developed to support financial literacy and user interaction in major Zambian languages such as Bemba, Nyanja, and Lozi. This promotes inclusivity and effective communication with end users.
- **REQ-05**: It is expected that biometric and multi-factor authentication mechanisms are incorporated to ensure secure and user-friendly access to the platform, reducing the risk of unauthorized access and identity theft.
- **REQ-06**: It is critical that all system components and data practices comply fully with Zambia's Data Protection Act (2021) and regulations established by the Zambia Information and Communications Technology Authority (ZICTA, 2023).
- **REQ-07**: It is critical that the infrastructure supporting the solution is scalable and low-latency, leveraging edge AI capabilities where applicable. This ensures real-time responsiveness and supports national-scale deployment.
- **REQ-08**: It is expected that there is a continuous mechanism for performance monitoring and impact evaluation, particularly in terms of financial inclusion metrics and the incidence of fraud. This supports ongoing improvement and accountability.

## 4 Sequence Diagram




## 5 References

[1] Alliance for Financial Inclusion. (2021) Smart Campaign's Client Protection Principles. Available at: https://www.afi-global.org (Accessed: 24 April 2025).

[2] Bank of Zambia. (2023) National Payment Systems Directives on Electronic Money Issuance, 2023. Lusaka: Bank of Zambia. Available at: https://www.boz.zm/Directive202307ElectronicMoneyIssuance2023.pdf (Accessed: 24 April 2025).

[3] International Telecommunication Union (ITU). (2023) 6G Framework for Developing Economies. Geneva: ITU.

[4] Masakhane. (2023) Decolonizing NLP for African Languages: Challenges, Opportunities, and Community Approaches. Presented at the Conference on Empirical Methods in Natural Language Processing (EMNLP) 2023. Available at: https://www.masakhane.io (Accessed: 24 April 2025).

[5] World Bank. (2020) AI in Credit Scoring: Opportunities for Financial Inclusion. Washington, DC: World Bank Group.

[6] Zambia Information and Communications Technology Authority (ZICTA). (2023) Annual Report on Cybersecurity Threats. Lusaka: ZICTA.

# Use Case 6: AI for Fraud Detection in NGO Operations: Bringing Trust with Data-Driven Governance

**Organization**: In

**Country:** India

**Contact Person(s)**:

Shivang Rattan, Independent Researcher, rattans123@gmail.com

Dr. Swapnil Morande, Independent Researcher, swapnil.morande@unina.it

## 1 Use Case Summary Table

| Item | Details |
|---|---|
| Category | Finance |
| Problem Addressed | Misuse of funds and fraudulent activities by NGOs, leading to inefficiencies in government and CSR funding. |
| Key Aspects of Solution | Three-tier AI system for proactive fraud detection:<br>1. AI-Based NGO Registration & Verification to detect duplicates and verify documents,<br>2. Real-Time Financial & Compliance Monitoring to flag suspicious activities<br>3. AI-Assisted Audits & Investigations to prioritize high-risk cases. |
| *Technology Keywords* | Registration & Verification: OCR, ML (duplicate detection), graph analytics.<br>Financial & Compliance Monitoring: ML (anomaly detection), real-time monitoring, risk scoring.<br>Audits & Investigations: NLP (sentiment analysis), geospatial analysis, satellite data. |
| *Data Availability* | Public |
| *Metadata (Type of Data)* | Text, images |

(continued)

| Item | Details |
|---|---|
| Model Training and Fine-Tuning | Anomaly Detection: Train models on NGO financial data to spot unusual transactions. Fine-tune with new data to catch evolving fraud.<br>Graph Analytics: Build networks of NGOs and related entities to find suspicious connections (e.g., duplicate registrations). Adjust detection parameters as needed.<br>Sentiment Analysis: Train NLP models on text data (social media, news) to detect negative sentiment or fraud allegations. Update with new text for better accuracy.<br>Geospatial & Satellite Data Analysis: Correlate reported project locations with satellite images to verify existence. Refine image analysis for better results.<br>OCR: Extract text from documents for digital analysis and cross-checking. |
| *Testbeds or Pilot Deployments* | Deployment is planned for Q3 2025 with select NGOs registered on NGO Darpan [1], focusing on financial transaction monitoring in Maharashtra. Proof of Concept (PoC) using synthetic data:<br>PoC Description: A PoC will use synthetic NGO financial data (e.g., 1,000 transactions across 50 NGOs) and registration records to test anomaly detection and duplicate registration algorithms. The PoC will simulate fraud scenarios (e.g., duplicate Permanent Account Numbers (PANs), unusual fund transfers) and evaluate model accuracy.<br>Dataset: Generate synthetic data mimicking NGO Darpan's structure (financial transactions, registration details).<br>Expected Outcome: Achieve >85% accuracy in detecting fraudulent transactions. |

# 2 Use Case Description

## 2.1. Description

In India, Non-Governmental Organizations (NGOs) play a crucial role in social development. However, the sector is vulnerable to various forms of fraud, including misuse of funds, duplicate registrations, and the operation of shell NGOs solely to siphon government and Corporate Social Responsibility (CSR) funds. These fraudulent activities undermine the effectiveness of genuine NGOs, lead to significant financial losses, and erode public trust. Traditional manual methods of scrutiny are often time-consuming, resourceintensive, and prone to human error, making it difficult to effectively identify and prevent sophisticated fraud.

Objectives: This use case aims to leverage the power of Artificial Intelligence to significantly enhance the detection and prevention of fraudulent activities within the NGO sector. The primary objectives are to:

- Automate the identification of financial irregularities and suspicious patterns in NGO transactions.
- Ensure greater compliance with regulatory frameworks such as the FCRA, CSR guidelines, and tax laws.
- Improve the transparency and accountability of fund allocation and utilization within the NGO ecosystem.
- Enable early detection of potential fraud, allowing for timely intervention and minimizing financial losses.

Technological Approach: This use case implements a multi-layered AI-driven system integrating several cutting-edge technologies:

- Machine Learning (ML): Algorithms will be trained on historical financial data to identify anomalies indicative of fraud, such as unusual transaction volumes or transfers to unverified entities.
- Graph Analytics: By building a network of relationships between NGOs, their directors, addresses, and financial transactions, this approach will detect interconnected fraudulent entities like duplicate or shell NGOs.
- Geospatial and Satellite Data Analysis: Satellite imagery and location data will be used to validate the existence and operation of claimed projects on the ground, crossreferencing NGO reports with real-world evidence.
- Natural Language Processing (NLP) and Sentiment Analysis: AI will analyse social media, news articles, and whistleblower reports to identify public complaints and allegations of financial mismanagement or corruption related to NGOs.
- Optical Character Recognition (OCR): This will automate the extraction of information from scanned documents during registration and compliance checks, enabling efficient data comparison and anomaly detection.

Expected Impact and Benefits: The implementation of this AI-driven fraud detection system is expected to yield significant benefits:

The implementation of the AI-driven fraud detection system for NGOs is anticipated to deliver transformative outcomes, enhancing efficiency, transparency, and trust in the NGO sector. Key impacts and benefits include:

- Significant Reduction in Fraudulent Activities: The system is projected to reduce fraudulent transactions by 30% within 12 months of pilot deployment by proactively detecting and deterring misuse of funds, duplicate registrations, and shell NGOs.
- Enhanced Operational Efficiency: Automation of registration verification, financial monitoring, and audit prioritization is expected to decrease audit processing time by 50%, reducing manual workload and enabling faster identification of suspicious activities.
- Real-Time Detection and Intervention: Continuous monitoring will enable early detection of irregularities, with 85% accuracy in risk scoring for high-risk NGOs, facilitating timely investigations and preventing significant financial losses.
- Robust Project Validation: Geospatial and satellite data analysis will validate 90% of reported project locations, ensuring funds are utilized for legitimate initiatives and minimizing discrepancies between reported and actual activities.

- Increased Transparency and Accountability: Real-time tracking of financial transactions and compliance, coupled with accessible audit trails, will ensure NGOs are held accountable, fostering a transparent ecosystem for government and CSR funding.
- Data-Driven Decision Making: AI-generated insights and risk scores will prioritize high-risk cases, enabling government bodies and donors to allocate resources effectively and focus audits on the top 5% of high-risk NGOs.
- Strengthened Public and Donor Confidence: By reducing fraud and enhancing transparency, the system will boost trust in the NGO sector, increasing donor willingness to fund legitimate social development initiatives by an estimated 20% over two years.

**Ethical and Privacy Considerations**:

- Data Privacy: All data will be encrypted and comply with India's Personal Data Protection Act, 2023. Access restricted to authorized personnel via secure APIs.
- Ethical AI: Models will be audited for bias (e.g., avoiding false positives for small NGOs). Whistleblower data will be anonymized.
- Transparency: Audit trails for AI decisions will be maintained for accountability.

**Use Case Status:** The use case is part of a larger research project.

### Partners

Partner: Indian Space Research Organisation (ISRO) ,Supplies satellite imagery for geospatial validation of NGO projects.

## 2.2 Benefits of use case

**Significant Reduction in Fraudulent Activities**: The system is projected to reduce fraudulent transactions by 30% within 12 months of pilot deployment by proactively detecting and deterring misuse of funds, duplicate registrations, and shell NGOs.

**Enhanced Operational Efficiency**: Automation of registration verification, financial monitoring, and audit prioritization is expected to decrease audit processing time by 50%, reducing manual workload and enabling faster identification of suspicious activities.

**Real-Time Detection and Intervention**: Continuous monitoring will enable early detection of irregularities, with 85% accuracy in risk scoring for high-risk NGOs, facilitating timely investigations and preventing significant financial losses.

**Robust Project Validation**: Geospatial and satellite data analysis will validate 90% of reported project locations, ensuring funds are utilized for legitimate initiatives and minimizing discrepancies between reported and actual activities.

**Increased Transparency and Accountability**: Real-time tracking of financial transactions and compliance, coupled with accessible audit trails, will ensure NGOs are held accountable, fostering a transparent ecosystem for government and CSR funding.

**Data-Driven Decision Making**: AI-generated insights and risk scores will prioritize high-risk cases, enabling government bodies and donors to allocate resources effectively and focus audits on the top 5% of high-risk NGOs.

**Strengthened Public and Donor Confidence**: By reducing fraud and enhancing transparency, the system will boost trust in the NGO sector, increasing donor willingness to fund legitimate social development initiatives by an estimated 20% over two years.

## 2.3 Future Work

The future work for this AI-driven fraud detection use case in NGO operations involves several key areas aimed at enhancing its effectiveness, scope, and real-world impact. The immediate next steps involve data collection and proof of concept development, with an initial phase focused on gathering relevant datasets from the identified sources (NGO databases, financial transaction records, regulatory bodies, geospatial data providers, social media platforms, news outlets) and building a functional prototype of the AI system. The proof of concept will involve testing the core AI/ML models for anomaly detection, duplicate identification, and initial compliance checks on a limited dataset to validate their feasibility and performance. Beyond this initial phase, future work will concentrate on:

- Advanced Model Development and Refinement: This includes exploring more sophisticated AI/ML techniques beyond initial anomaly detection, such as predictive modelling to anticipate potential fraud, and developing more nuanced algorithms for identifying complex patterns indicative of shell NGOs or sophisticated financial manipulation. Continuous fine-tuning of existing models based on new data and feedback from investigators will be crucial to maintain accuracy and adapt to evolving fraud tactics.
- Enhanced Data Integration and Interoperability: Expanding the system's ability to seamlessly integrate with a wider range of data sources, including more granular financial data, project implementation reports, and beneficiary databases, will provide a more holistic view of NGO operations and improve detection accuracy. This may involve developing robust APIs and data pipelines for efficient data exchange.
- Real-time Monitoring and Alerting Capabilities: Enhancing the system to provide more sophisticated real-time monitoring and alerting mechanisms will be a priority. This could involve developing customizable alert thresholds, integrating with existing notification systems, and providing more contextual information with alerts to aid in faster analysis.
- Visualization and Reporting Tools: Developing user-friendly dashboards and reporting tools will be essential for analysts and decision-makers to effectively interpret the AIgenerated insights, visualize potential fraud networks, and generate comprehensive reports for investigations and policy recommendations.
- Ground-Level Validation Enhancement: Improving the ground-level validation capabilities by incorporating more diverse data sources beyond satellite imagery, such as drone footage analysis, integration with on-the-ground reporting mechanisms, and potentially even citizen-sourced data (with appropriate safeguards), could provide more robust verification of project existence and impact.

**Project Timeline**

| Timeline | Milestone |
|---|---|
| Q2 2025 | Data collection and PoC development using synthetic data. |
| Q3 2025 | Pilot deployment in Maharashtra with 50 NGOs. |
| Q4 2025 | Model refinement and integration with additional data sources (e.g., MCA, ISRO). |
| 2026 | Nationwide rollout and dashboard development. |

**Sample Dashboard Design & Report Format**

**Proposed Dashboard Design**:

### Overview Tab:

- KPI Metrics: Number of NGOs monitored, fraud alerts generated, projects validated.
- Risk Score Distribution: Bar chart of NGOs by risk level (low, medium, high).

**Fraud Detection Tab**:

- Anomaly Alerts: Table listing flagged transactions (NGO name, amount, date, risk score).
- Network Visualization: Graph showing connections between suspicious NGOs/directors.

### Geospatial Validation Tab:

- Map View: Overlay of reported project locations with satellite imagery.
- Validation Status: Pie chart (verified, unverified, flagged).

### Sentiment Analysis Tab:

- Sentiment Trends: Line chart of public sentiment (positive/negative) over time.
- Keyword Cloud: Top fraud-related terms from social media/news.

**Technology**: Built using Plotly Dash for interactivity.

### NGO Fraud Detection Report

- Sections:

1. Summary: Total NGOs analyzed, fraud alerts, and high-risk cases.
2. Financial Anomalies: List of flagged transactions with details (amount, recipient, risk score).
3. Registration Issues: Duplicate registrations or unverified documents.
4. Project Validation: Discrepancies between reported and actual project locations.
5. Public Sentiment: Summary of negative sentiment and key allegations.
6. Recommendations: Prioritize audits for top 5 high-risk NGOs; investigate specific transactions.

- Format: PDF exportable from dashboard with NIC branding.

## 3 Use Case Requirements

- **REQ-01**: The system shall ingest and analyse financial data to detect anomalous transactions and patterns indicative of fraud.
- **REQ-02**: The system shall identify potentially fraudulent NGOs by detecting duplicate registrations and suspicious operational or financial profiles.
- **REQ-03**: The system shall monitor NGO compliance with key regulatory requirements, including reporting obligations.
- **REQ-04**: The system shall leverage location-based data and external imagery to validate the existence and activities of claimed projects.

- **REQ-05**: The system shall analyze textual data from various sources (social media, news, reports) to identify potential fraud indicators and public sentiment.
- **REQ-06**: The system shall generate alerts and provide risk scoring to prioritize potential fraud cases for review.
- **REQ-07**: The system shall integrate with relevant government databases for data verification and cross-referencing.
- **REQ-08**: The system shall ensure secure data handling and provide a user interface for authorized access and investigation.
- **REQ-09**: The underlying AI/ML models shall be adaptable and updated to maintain detection accuracy.

## 3.1 Separate AI Models and Infrastructure

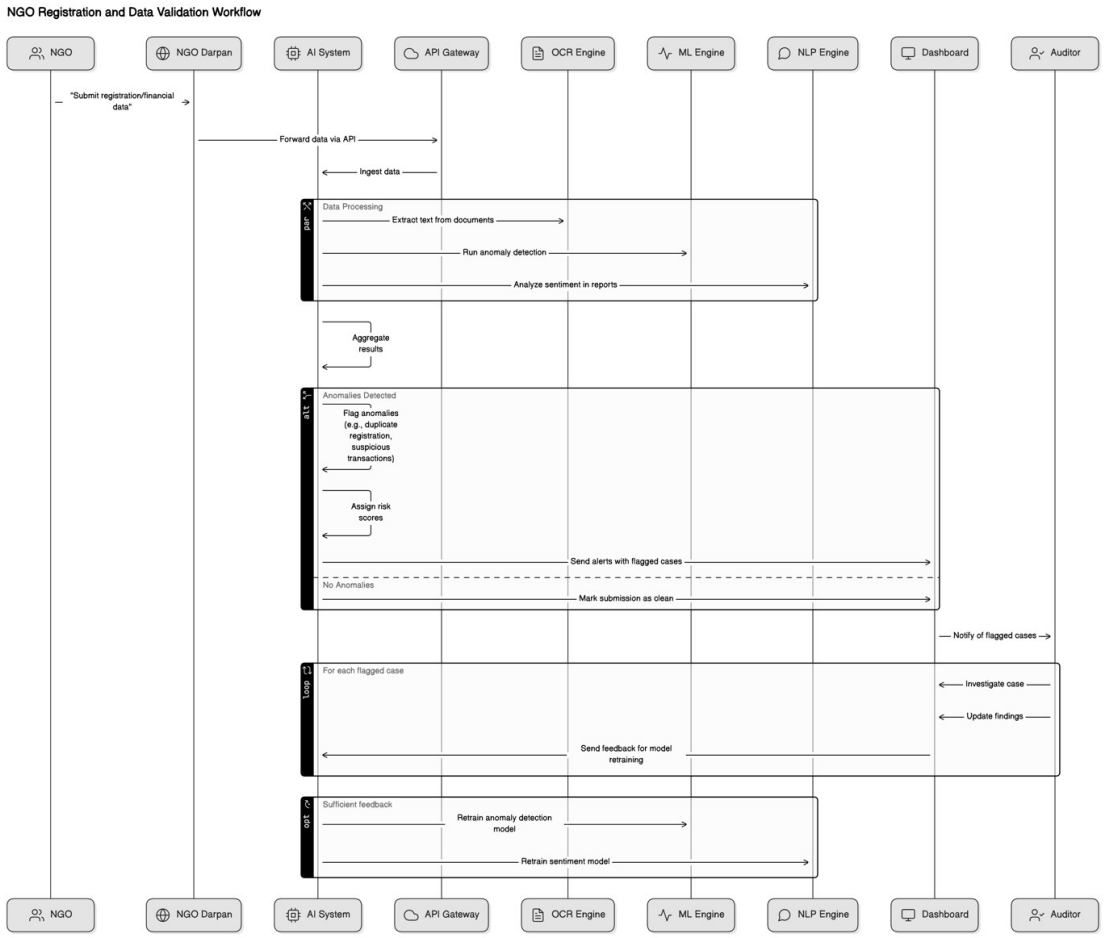
### AI Models:

- Anomaly Detection: Trained on financial data using scikit-learn or TensorFlow for detecting unusual transactions.
- Graph Analytics: Uses NetworkX to identify duplicate registrations via entity networks.
- NLP/Sentiment Analysis: Fine-tuned BERT models for analyzing social media/news.
- Geospatial Analysis: Custom algorithms to correlate satellite imagery with project locations.
- OCR: Tesseract for document text extraction.

### Infrastructure:

- Data Sources:
- Primary: NGO Darpan (ngodarpan.gov.in) for registration and financial data.
- Secondary: ISRO for satellite imagery, MCA for CSR data, social media APIs (e.g., Twitter/X).
- APIs:
- RESTful APIs for real-time data ingestion from NGO Darpan and MCA.
- ISRO Bhuvan API for geospatial data.
- Twitter/X API for sentiment analysis.
- Data Collection Protocols:
- Secure HTTPS for data transfer.
- Batch processing for historical data; streaming for real-time updates.
- SOPs:
- Data Ingestion: Validate data formats (CSV, JSON) and log errors.
- Model Training: Retrain models quarterly with new data; validate with 80/20 train-test split.
- Alert Handling: Auditors review high-risk alerts within 48 hours; escalate critical cases to NIC leadership.
- Privacy **Compliance**: Anonymize whistleblower data; encrypt financial records per India's PDP Act.

# 4 Sequence Diagram




# 5 References

[1] NGO Darpan Portal. *NGO Darpan*. Accessed June 27, 2025. https://ngodarpan.gov.in.

[2] Ministry of Electronics and Information Technology. *The Digital Personal Data Protection Act (2023)*. Accessed June 27, 2025. https://www.meity.gov.in/data-protection-act2023.

[3] Indian Space Research Organisation. *ISRO Bhuvan*. Accessed June 27, 2025. https://bhuvan.nrsc.gov.in.

# Use Case 7: Strengthening Financial Integrity and International Trust: UAE's FATCA/CRS Framework

UNITED ARAB EMIRATES
MINISTRY OF FINANCE

الإمارات العربية المتحدة
وزارة المالية

**Country**: United Arab Emirates

**Organization**: Ministry of Finance

**Contact Person(s)**:

**Hana Akram,** hhakram@mof.gov.ae

**Asma Alzarooni,** amalzarooni@mof.gov.ae

## 1 Use Case Summary Table

| Item | Details |
|---|---|
| Category | Finance |
| Problem Addressed | Lack of transparency and standardization in cross-border financial reporting enables tax evasion and limiting the ability of authorities to detect and act on non-compliance. |
| Key Aspects of Solution | Implementation of a centralized digital system enabling secure registration, risk assessment, data validation, and international information exchange.<br>The system enhances compliance monitoring through automated audit tools and facilitates data sharing with over 80 jurisdictions, ensuring alignment with FATCA and CRS standards. |
| Technology Keywords | Data Transmission: International Data Exchange Service (IDES) for FATCA, Common Transmission System (CTS) for CRS<br>Identity / Security: Login through UAEPASS (UAE's national digital identity<br>Data packaging: XML Data Packaging<br>Data Quality Management: Input data validation rules (including TIN format validation algorithm)<br>Compliance: Risk assessment questionnaire and audit assessment |
| Data Availability | Data submitted by Reporting Financial Institutions (RFI) is private and securely stored within the UAE Ministry of Finance FATCA/CRS System. Access is restricted to authorized personnel only. Each Regulatory Authority can access data only for the RFIs under its remit and cannot view data submitted to other authorities. No public repository is available due to confidentiality requirements under FATCA and CRS agreements. |
| Metadata (Type of Data) | Structured data (e.g., account information, financial details, risk ratings), tabular data (Excel/CSV) (e.g. data submission, reports generated), and XML files for standardized reporting and transmission. |

(continued)

| Item | Details |
|---|---|
| Model Training and Fine-Tuning | Model training is not applicable. The current system is rule-based and does not involve model training or fine-tuning. Future enhancements may explore the integration of AI for predictive risk analysis or anomaly detection.<br>However, as part of the UAE Ministry of Finance's broader digital transformation efforts, the official MoF website now includes an AI-powered conversational assistant to support users with inquiries related to the FATCA/CRS system. This feature, developed under the Zero Bureaucracy initiative, provides real-time guidance and improves user experience, aligning with the UAE's vision for smart government services. |
| Testbeds or Pilot Deployments | The FATCA/CRS platform is a fully deployed national solution in the UAE. It is used by over 1,400 Reporting Financial Institutions and supports the secure exchange of tax data with 80+ jurisdictions. The system is integrated with international transmission platforms, including IDES (for FATCA/IRS) and CTS (for CRS/OECD), and is actively used for annual reporting cycles. |

# 2 Use Case Description

## 2.1 Description

The lack of transparency in cross-border financial activities poses significant challenges to tax compliance, revenue collection, and international fiscal cooperation. To address these concerns, global frameworks such as the U.S. Foreign Account Tax Compliance Act (FATCA) and the OECD's Common Reporting Standard (CRS) were introduced to enable the automatic exchange of financial information between jurisdictions [1].

The UAE has demonstrated a strong commitment to these standards. FATCA has been in effect in the UAE since 1 July 2014, requiring financial institutions to report information about U.S. account holders to the U.S. IRS. Similarly, the UAE adopted the CRS on 1 January 2017, obligating it to collect and exchange data on tax residents of partner jurisdictions through the OECD network [1].

In response, the UAE implemented a digital FATCA/CRS reporting system that enables financial institutions to register, submit data, undergo risk assessment, and comply with reporting obligations. The system also incorporates secure data exchange mechanisms with partner jurisdictions via IRS's IDES and OECD's CTS platforms.

Recognizing the evolving nature of global standards, the UAE has also implemented robust audit and assessment programs to ensure continued alignment and high-quality data reporting. These efforts have positioned the UAE as a proactive player in international tax cooperation and transparency.

The system is designed to be adaptive and scalable, supporting future enhancements such as the implementation of CRS 2.0 requirements. Additionally, the same infrastructure is being leveraged to support the CARF (Crypto-Asset Reporting Framework), promoting a unified and efficient approach to regulatory compliance in the financial sector.

Integration with Key Technologies and APIs

The FATCA/CRS system leverages a range of key technologies and APIs to ensure secure, seamless, and standards-compliant operations:

- UAE PASS is integrated for secure, passwordless user authentication, ensuring that only authorized users can access sensitive data and system functionalities [3].
- IDES (International Data Exchange Service), developed by the U.S. IRS, is used for the secure electronic transmission of FATCA data, enabling compliance with international reporting obligations.
- CTS (Common Transmission System), developed by the OECD, ensures the efficient and secure exchange of CRS data with partner jurisdictions [4].
- Advanced Data Validation mechanisms are embedded within the system to ensure the accuracy of critical identifiers such as Global Intermediary Identification Numbers (GIINs) and Tax Identification Numbers (TINs), minimizing reporting errors and enhancing data quality.

As part of the UAE Ministry of Finance's broader digital transformation efforts, the official MoF website now includes an AI-powered conversational assistant to support users with inquiries related to the FATCA/CRS system. This feature, developed under the Zero Bureaucracy initiative, provides real-time guidance and improves user experience, aligning with the UAE's vision for smart government services.

### Impact and Benefits

- Improved tax transparency and prevention of tax evasion by both individuals and entities[1][2].
- Strengthened international collaboration through successful data exchange with over 80 jurisdictions [3].
- Enhanced compliance with global tax regulations and reporting standards.
- Maintained confidentiality and security of financial data in accordance with OECD and IRS XML protocols.
- Strengthened the capacity of the Ministry of Finance (MoF) and Regulatory Authorities (RAs) to effectively implement and oversee FATCA/CRS obligations.
- Reinforced the UAE's reputation as a committed and responsible member of the OECD Global Forum and the IRS international network.

## 2.2 Benefits of the use case

### Decent Work and Economic Growth

Promote development-oriented policies that support productive activities, decent job creation, entrepreneurship, and encourage the formalization and growth of micro-, small- and medium-sized enterprises.

Link to Project: By improving tax transparency and reducing illicit financial flows, the FATCA/CRS project helps create a fairer economic environment that supports sustainable growth. Transparent financial systems build investor confidence, encourage entrepreneurship, and ensure that revenues are available for economic and social development, contributing to inclusive employment and economic opportunity.

### Reduced Inequalities

Improve the regulation and monitoring of global financial markets and institutions and strengthen the implementation of such regulations.

Link to Project: The project contributes directly to reducing inequality between and within countries by ensuring that financial assets are disclosed and taxed fairly across jurisdictions. By holding individuals and institutions accountable for cross-border financial activities, the system prevents tax evasion that often exacerbates income inequality.

### Partnerships for the Goals

Strengthen domestic resource mobilization, including through international support to developing countries to improve domestic capacity for tax and other revenue collection. Enhance capacity-building support to developing countries, including for data collection and monitoring.

Link to Project: FATCA/CRS implementation is built on international cooperation and data exchange with 80+ jurisdictions. This strengthens partnerships and supports global efforts to mobilize resources through tax compliance. It also enhances the UAE's domestic capacity through improved audit tools, data validation processes, and inter-agency coordination, contributing to stronger institutions and governance.

### Partners

- UAE Ministry of Finance (role: owner of the product)
- Regulatory Authorities (RAs): (role: users of the product)
- Central Bank of the United Arab Emirates
- Dubai International Financial Centre
- Abu Dhabi Global Markets
- Securities and Commodities Authority
- Federal Tax Authority
- Financial Institutions (e.g., banks, insurance firms, fund management entities): (role: users of the product)
- International Bodies: (role: stakeholder entities, to whom data is exchanged)
- Organization for Economic Cooperation & Development (OECD) for the Common Reporting Standard (CRS)
- Internal Revenue Service (IRS) for the Foreign Account Tax Compliance Act (FATCA)

## 2.3 Future Work

**I. Planned system releases**

Audit

- Regulatory Authorities conduct audits by issuing questionnaires to Reporting Financial Institutions to assess non-compliance, including reviewing sampled documentation.
- The audit process includes selecting specific account samples, requesting and reviewing documentation, and evaluating findings.

- An annual report is prepared by the Regulatory Authorities summarizing audit activities and year-on-year trends.

Enhanced Reporting Templates

- Enhanced reporting templates enable the collection of additional data fields from Reporting Financial Institutions to support TIN improvement and enhanced input data validations.
- Enhances data validation checks to improve data quality, facilitate data review, and strengthen compliance with reporting requirements.

Data Review

- Data review enables Regulatory Authorities and the MoF Competent Authority to review data post-submission by Financial Institutions.
- This reporting functionality enables the MoF Competent Authority to perform targeted outreach to the Reporting Financial Institutions to remediate any data issues if observed.
- This allows improvement of data quality prior to transmission to the partner jurisdictions.

### II. Additional Resources Needed

To ensure continued alignment with evolving international standards and to enhance system performance and scalability, the following additional resources are identified:

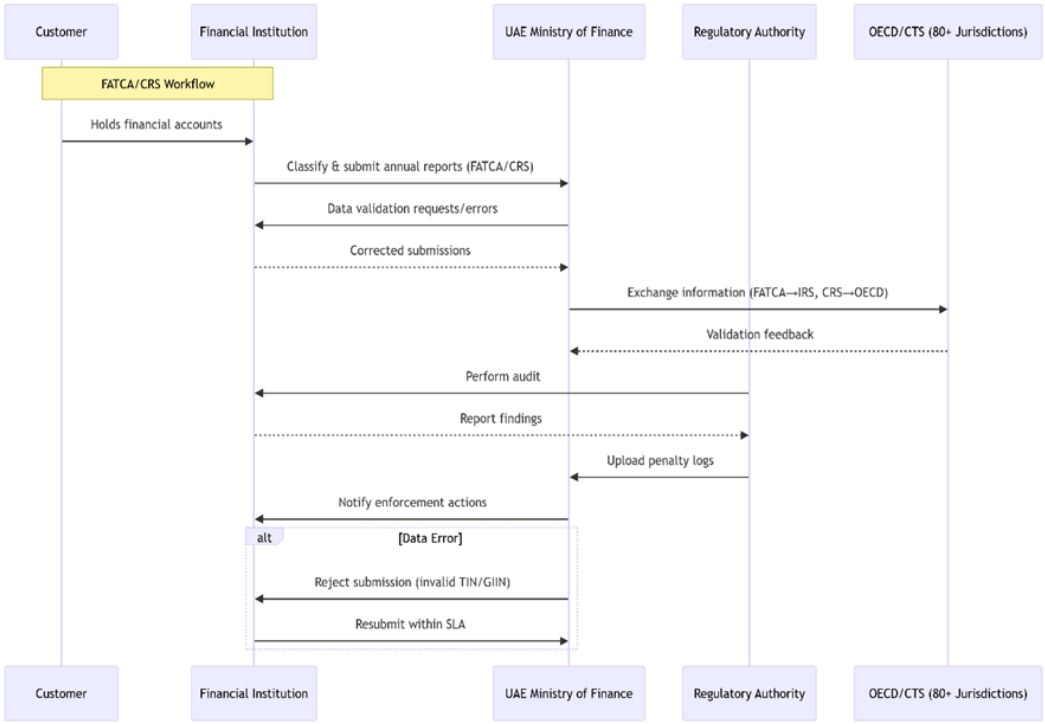
- Technology Upgrades: Enhancements to the current platform architecture may be required to support the integration of new frameworks such as CRS 2.0 and the CARF (Crypto-Asset Reporting Framework). This includes upgrading the system's API infrastructure.
- Advanced Validation Capabilities: The incorporation of more sophisticated validation rules will be necessary to meet upcoming requirements under CRS 2.0 and CARF. This may include rule-based logic for new data fields and cross-validation against global reference datasets.
- Data Collection Expansion: To comply with future reporting obligations, the system will need to capture additional data points and expanded taxpayer information under CARF.
- Training and Capacity Building: Continued investment in training for both Regulatory Authorities and financial institutions will be required to ensure proper understanding and adoption of new reporting standards and system functionalities.

## 3 Use Case Requirements

- **REQ-01**: It is critical that the system enables Reporting Financial Institutions (RFIs) to register their entities through a secure portal. It supports a maker-checker review workflow and is integrated with UAE PASS for seamless login and authentication.
- **REQ-02**: It is critical that the platform supports Excel-based data submissions from RFIs in accordance with reporting obligations. Built-in input validation ensures erroneous or incomplete data is automatically flagged prior to submission
- **REQ-03**: It is expected that the system performs automated computation of risk ratings based on responses to a standardized risk assessment questionnaire. These ratings are accessible to Regulatory Authorities for use in audit planning and compliance monitoring.
- **REQ-04**: It is critical that submitted data is converted into XML format for secure exchange with international partner jurisdictions. The Ministry of Finance (Competent Authority) manages this exchange and can initiate data corrections when needed.

- **REQ-05**: It is expected that the system enables Regulatory Authorities to upload a log of penalties applied as part of enforcement actions. It is expected that Competent Authority users have access to view all enforcement actions logged by the Regulatory Authorities.
- **REQ-06**: It is critical that the system applies strict validations on key fields such as Tax Identification Numbers (TINs), Global Intermediary Identification Numbers (GIINs), and account details. All mandatory fields are checked for completeness and conformance with OECD and IRS XML schema standards before data is accepted into the system. It is critical that the TIN validation algorithm compares TIN formats against 120+ jurisdictions as per specifications published by OECD. It is critical that the system packages XMLs by tax residency of account holders and transmits them to the respective partner jurisdictions.

## 4 Sequence Diagram




## 5 References

[1] OECD, *Standard for Automatic Exchange of Financial Account Information in Tax Matters: Second Edition*. [Online]. Available: https://www.oecd.org/en/publications/standard-for-automatic-exchange-of-financial-account-information-in-tax-matters-second-edition_9789264267992-en.html?utm_source=chatgpt.com

[2] Internal Revenue Service (IRS), "Foreign Account Tax Compliance Act (FATCA)," [Online]. Available: https://www.irs.gov/businesses/corporations/foreign-account-tax-compliance-act-fatca

[3] UAE Ministry of Finance, "FATCA and CRS," [Online]. Available: https://mof.gov.ae/fatca-and-crs/

[4] UAE Ministry of Finance, *FATCA-CRS System RFI User Guide v2.3*, Feb. 2025. [Online]. Available: https://mof.gov.ae/wp-content/uploads/2025/02/FATCA-CRS-System_RFI-User-Guide_v2.3.pdf

# Use Case 8 Sustainable Financing Partnership

**Organization:** Mohammed Bin Rashid for Housing Establishment

**Country:** United Arab Emirates(UAE)

**Contact Person(s):**

Aseel Ayoub , aayoub@mbrhe.ae

Fatma Al ustad, falustad@mbrhe.ae

## 1 Use Case Summary Table

| Item | Details |
|---|---|
| Category | Finance |
| Problem Addressed | Mohammed Bin Rashid Housing Establishment (MBRHE) was constrained by limited annual budgets and increasing demand for housing loans. Traditional internal financing models limited the number of beneficiaries, caused service delays, and placed significant pressure on internal human resources. |
| Key Aspects of Solution | Outsourcing of housing loan funding to Emirates Islamic Bank (EIB) through a sustainable financing agreement.<br>MBRHE guarantees interest payments while the bank disburses the principal loan amount.<br>AI models infer customer eligibility and priority based on credit, income, and social criteria.<br>Real-time integration with the bank via secure application program interface(APIs).<br>Loan service is accessible to customers through the UAE PASS-enabled digital interface. |
| Technology Keywords | AI, predictive analytics, APIs, e-government, digital identity, bank integration, outsourcing |
| Data Availability | Public |
| Metadata (Type of Data) | Text |
| Model Training and Fine-Tuning | AI models are trained on historical MBRHE loan approval data, augmented with credit bureau data. Techniques include logistic regression, decision trees, and ensemble methods for credit risk prediction and social vulnerability scoring. |
| Testbeds or Pilot Deployments | Piloted at 20 customer centres in Q2 2024. Full deployment across Dubai completed in Q1 2025. Internally documented by MBRHE. |

# 2 Use Case Description

## 2.1 Description

MBRHE's commitment to providing housing for Emirati citizens was increasingly challenged by growing demand, fixed budgets, and resource-intensive internal loan processing. The Sustainable Financing Partnership use case introduced a new operating model where loans are financed directly by Emirates Islamic Bank, with MBRHE subsidizing interest. This model allows MBRHE to provide loans without a funding cap and reduce internal operational burden.

Through AI-based credit and eligibility models, citizens are automatically evaluated for loan disbursement, removing bottlenecks in manual approval. Real-time integration between MBRHE and EIB enables instant decision, tracking, and fund disbursement. Customers access services through a digital platform using UAE PASS, allowing fully remote application, approval, and servicing.

**Use Case Status:** Fully implemented (Q1 2025)

### Partners:

Mohammed Bin Rashid Housing Establishment (MBRHE), Emirates Islamic Bank (EIB) , Dubai Digital Authority, UAE Credit Bureau

## 2.2 Benefits of the use case

- Our initiative offers significant benefits by accelerating access to affordable finance through targeted subsidies, which directly reduces housing poverty among vulnerable groups. By leveraging advanced technologies such as AI, APIs, and cloud microservices, we modernize the public housing finance infrastructure, making it more efficient and accessible. This technological integration not only enhances urban resilience by expanding inclusive housing finance options but also reduces systemic bottlenecks that often hinder progress. Additionally, our approach fosters greater trust in public institutions by ensuring transparent and efficient digital service delivery, thereby strengthening the overall governance and effectiveness of public programs.
- The expected impact of this initiative is substantial. It enables unlimited loan disbursement based on eligibility, ensuring that those in need have access to the financial support they require. With a customer satisfaction index of 98.4%, the program has proven to be highly effective in meeting the needs of its beneficiaries. Over 2,100 loans were issued in the first six months alone, demonstrating the program's immediate and tangible impact. The initiative also brings significant financial savings, with an estimated UAE Dirham(AED) 246 million saved annually through operational and budget efficiencies. Furthermore, by reducing internal workload by 30%, the program enables staff to be redeployed to strategic initiatives, enhancing overall organizational effectiveness and capacity.

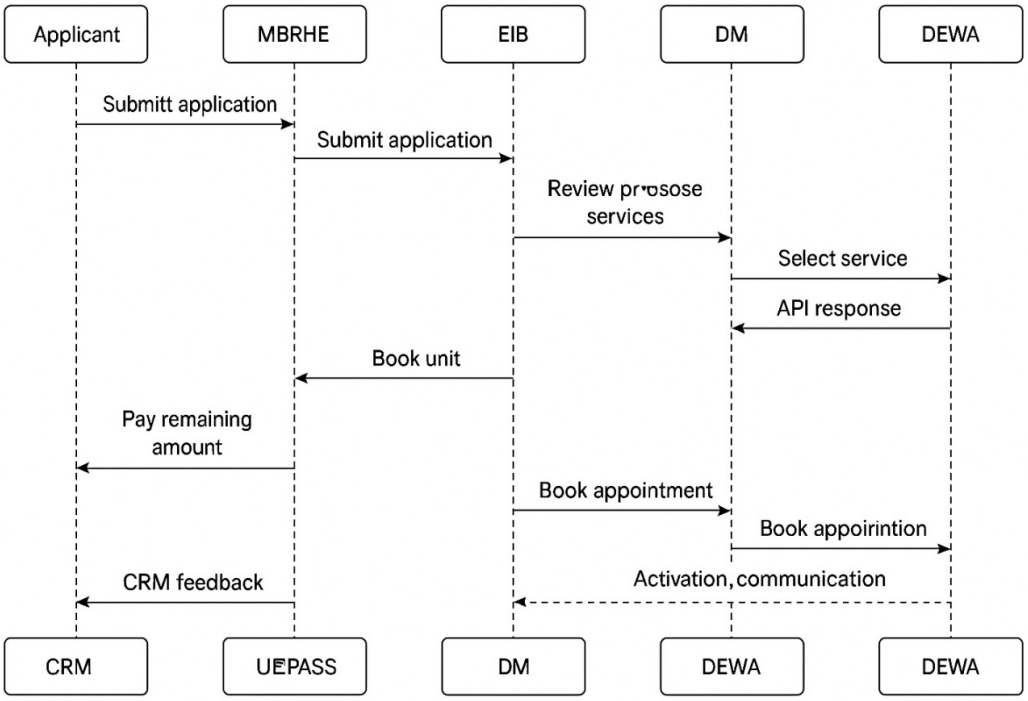
## 2.3 Future Work

- Expand partnership to include other banks to ensure competitive rates and wider reach. Enhance AI models with alternative data sources (e.g., utility payments, behavioral data). Develop customer-facing mobile app for status tracking and support. Publish a whitepaper on AI models used to enhance transparency and collaboration with global peers. Explore expansion to non-traditional housing segments (e.g., renovations, green homes)

## 3 Use Case Requirements

- **REQ-01**: AI engine for eligibility and credit scoring using real-time customer data
- **REQ-02**: Secure API integration between MBRHE and EIB for data exchange and processing
- **REQ-03**: Cloud-based infrastructure to support scalability and high availability
- **REQ-04**: Secure digital user authentication via UAE PASS
- **REQ-05**: Feedback loop through Customer Relationship Management(CRM) for post-loan servicing and satisfaction tracking

## 4 Sequence Diagram




## 5 References

[1] Dubai Social Agenda 2033 – For alignment with long-term national goals.

[2] Dubai Digital Strategy – Supporting your digital integration, AI, and UAE PASS use.

[3] Dubai Paperless Strategy – Reinforces digital document workflows and system automation.

[4] Dubai Government Customer Happiness Index – For benchmarking MBRHE's No.1 ranking in customer satisfaction.

# Use Case 9: Receipt Guardian

Organization Details

**Organization**: Higher Colleges of Technology

**Country**: United Arab Emirates (Al Ain)

**Contact Person(s):**

Mohamed Elsersy (melsersy@hct.ac.ae)

Mariam Abdulwahab Alrayssi (H00412697@hct.ac.ae)

## 1 Use Case Summary Table

| Item | Details |
|---|---|
| Category | Accessibility, Financial Technology (FinTech), Family & Youth Financial Education |
| Problem Addressed | Manual expense tracking is inefficient, prone to errors, and lacks family-wide visibility |
| Key Aspects of Solution | • Mobile app with OCR<br>• Budget tracking and alerts<br>• Role-based sharing (parents/children)<br>• Security (AES-256, MFA) |
| Technology Keywords | OCR, Android, Firebase, AES-256, Multi-Factor Authentication (MFA), GitHub Actions, AWS, CI/CD |
| Data Availability | Private (scanned receipts by users, no public datasets mentioned) |
| Metadata (Type of Data) | Visual (receipt images), Text (extracted receipt info), User input (budget, notes) |
| Model Training & Fine-Tuning | OCR logic (pre-trained models adapted), error correction using user input suggestions |
| Testbeds / Pilot Deployments | Internal UAT and beta testing with >10 users, Android Studio emulator, family use cases |
| Code Repositories | Managed with Android Studio. |

# 2 Use Case Description

## 2.1 Description

Context & Background: Managing personal finances and expenses is challenging, especially for families. Parents often struggle to monitor children's spending, and manual receipt storage is outdated and time-consuming.

**Objectives:**

- Provide a modern, paperless solution for tracking expenses.
- Allow parents to monitor children's spending habits.
- Automate data capture through OCR.
- Offer secure, accessible, and user-friendly budget management.

**Technological Innovation**:

- Uses OCR for accurate data extraction from receipts.
- AES-256 encryption for financial data security.
- MFA and session management for safe access.
- Android mobile app integrated with Firebase and AWS.

**Expected Impact**:

- Promotes financial literacy among youth.
- Empowers families to track and manage household budgets.
- Encourages secure and responsible spending habits.

**Use Case Status**: Prototype complete, tested and validated. Beta deployed and pending public launch.

## 2.2 Benefits of use case

The app promotes financial literacy in children and supports mindful spending and budgeting habits. The app equips youth and families with tools to manage spending, encouraging education and sustainability.

- Development Team: Built and tested the mobile app.
- Test Users: Parents and children providing feedback during UAT.
- AWS / Firebase: Technology providers for backend hosting and data storage.

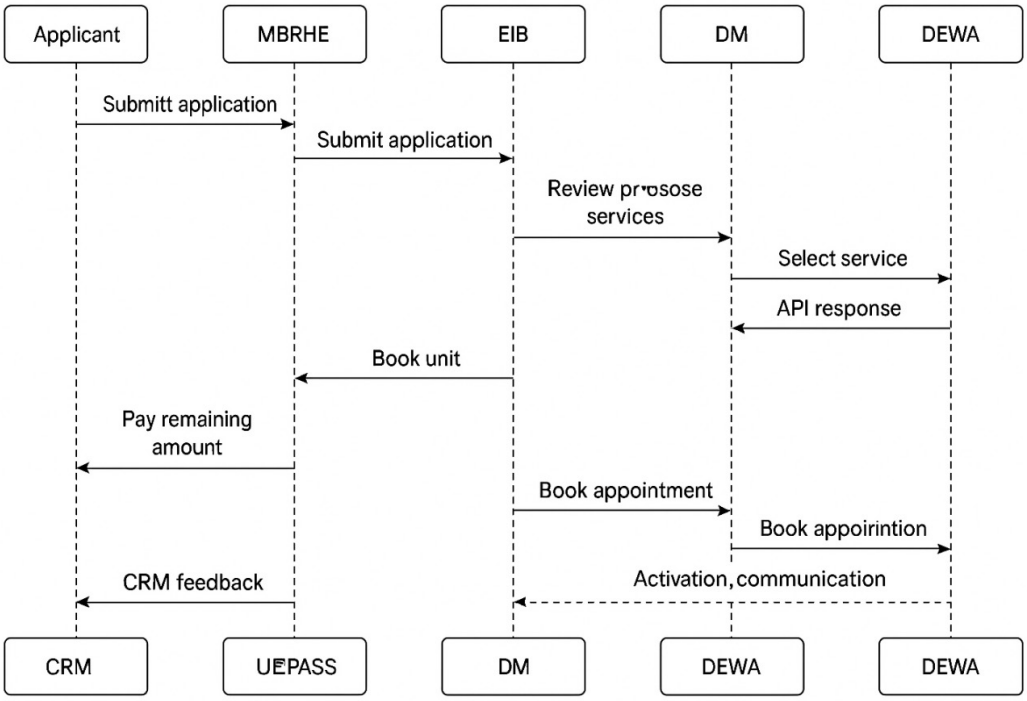
## 2.3 Future Work

- Improve OCR precision for multi-language receipts.
- Expand to iOS and web versions.
- Integrate with banking APIs for real-time syncing.
- Add gamified savings goals for children.
- Enhance visualizations using AI-based spending predictions.
- Seek institutional collaborations with schools or NGOs focused on youth financial literacy.

## 3 Use Case Requirements

- **REQ-01** App must extract text from various receipt formats using OCR.
- **REQ-02** Store user data securely with AES-256 encryption.
- **REQ-03** Authenticate users using MFA and session management.
- **REQ-04** Display categorized expenses and budgets clearly.
- **REQ-05** Support Android versions 10-12.
- **REQ-06** Handle blurry/misaligned receipts with correction options.
- **REQ-07** Deploy via CI/CD (GitHub Actions).
- **REQ-08** Enable role-based access (parent/child).

## 4 Sequence Diagram




Actors: User (Parent/Child), Mobile App, OCR Module, Firebase DB, AWS Storage

Flow:

1. User logs into app → Authenticates (MFA)
2. User uploads receipt → Image sent to OCR module
3. OCR extracts text → Parsed into expense data
4. Data stored in Firebase → Receipt image in AWS S3
5. Budget/expenses updated → Feedback shown on UI
6. Parent views child's expense summary → Notifications sent

## Expert Feedback and Proposed Solutions

1) Is it possible to add loans and credit cards?

Yes, it's possible.

Simple solutions:

- Let users manually add loan or credit card entries.
- Set reminders for payment due dates.
- Show remaining balance for each loan/card.

2) How can we ensure user privacy, especially for credit card information?

User privacy is very important.

Simple solutions:

- Don't store full card numbers (only last 4 digits if needed).
- Use strong encryption (AES-256 is already used – great!).
- Enable Multi-Factor Authentication (already included – well done).

3) What types of models are needed? (OCR, classification, recommendations, etc.)

Here are useful models:

| Model Type | Purpose |
|---|---|
| OCR | Read text from receipts (already done) |
| Image Processing | Improve photo quality (fix blur, angle) |
| Classification | Categorize receipts (e.g., food, travel) |
| Recommendation | Suggest budgets or savings goals |
| Prediction | Forecast future spending |

4) What is the current status of the project?

The application is largely working, but it needs some modifications. It is currently in the actual testing phase and is being developed.

- Internal testing (UAT) completed.
- Feedback is being collected to improve the app before public release.

5) Can the app connect with WhatsApp, WeChat, or Email?

Yes, and it's a very useful idea.

Simple solutions:

- Email: Auto-scan receipts from emails (e.g., Gmail).
- WhatsApp: Use WhatsApp Business API to import receipts.

## 5 References

[1] Smith, A. (2022). Optical Character Recognition: Best Practices for Receipt Digitization. TechJournal

[2] Deploying Mobile Apps with AWS Amplify. AWS

[3] Mobile Application Security Verification Standard (MASVS). OWASP

[4] Nielsen, J. (2021). Usability Testing Guidelines. NNG

[5] Automating Deployment with GitHub Actions. GitHub Docs.

# Use Case 10: Dynamic Threshold Adjustment for Credit Scoring: An Adaptive Optimization Approach for Subprime Lending

**Organization:** Vellore Institute of Technology, Chennai

**Country:** India

**Contact Person(s):**

Parth Khairnar, pskhairnar1024@gmail.com

Dr. Geetha S, geetha.s@vit.ac.in

## 1 Use Case Summary Table

| Item | Details |
|---|---|
| Category | Finance |
| Problem Addressed | Traditional static credit thresholds fail to account for individual behavior changes or economic shifts, resulting in unfair credit access denial. |
| Key Aspects of Solution | • Adaptive thresholding using Markov Chain Monte Carlo (MCMC)<br>• Predictive modeling with Light Gradient Boosting Machine(GBM)<br>• Hyper-parameter tuning via Bayesian Optimization<br>• Receiver Operating Characteristic - Area Under the Curve(ROC-AUC) metrics |
| Technology Keywords | LightGBM, MCMC, Bayesian Optimization, Hyperopt, Credit Risk Modeling, Fair AI, Python |
| Data Availability | Public [1] |
| Metadata (Type of Data) | Text |
| Model Training and Fine-Tuning | • LightGBM model trained with K-Fold Cross Validation<br>• MCMC for posterior sampling<br>• Bayesian Optimization for hyper-parameter tuning |
| Testbeds or Pilot Deployments | Initial testing in Google Colab with synthetic simulations Potential for sandbox testing with financial partners [2] |

## 2 Use Case Description

### 2.1 Description

Credit scoring has traditionally relied on logistic regression, decision trees, and more recent methods like random forests or neural networks to calculate an explicit probability threshold

for lending. These methods screen potential borrowers based on characteristics like payment history, debt amount, credit history age, credit mix employed, recent behavior, and credit utilization, each contributing to a credit score that reflects the creditworthiness of a person. However, more advanced methods are now suggested by recent studies: Leevy et al. [3] demonstrate that optimization of the decision threshold with a constraint TPR ≥ TNR can boost Area Under the Precision-Recall Curve (AUPRC) without resampling; Kamimura et al. [4] survey forty-six metaheuristic and Bayesian methods—like genetic algorithms or tree-structured Parzen estimators—to optimize credit scoring models; Kyeong & Shin [5] introduce a two-stage Bayesian logistic model to offer a trade-off between interpretability and performance; and the Khashei & Mirahmadi hybrid soft-computing model indicates improvements in cases of high uncertainty. Besides scoring, anomaly-detection techniques (like isolation forests, autoencoders) and fraud-detection programs introduce additional levels of risk regulation.

As a consequence of these findings, our use case implements an adaptive real-time decision-support system whose acceptance threshold it is continually adapting. We make use of MCMC (Metropolis–Hastings) to sample from the posterior distribution of threshold parameters and Bayesian Optimization (through Hyperopt) to hyperparameter-tune a LightGBM model.

In order to put these innovations into practice, we utilize the Home Credit Default Risk dataset, which comprises a chain of linked tables that provide an overall snapshot of each applicant's financial history and behavior. Demographic information, incomes, credit amounts, employment and education data, and housing and family status for each applicant are found in the application_train and application_test.csv datasets. Supplementing this data, the bureau.csv and bureau_balance.csv tables contain the applicant's external credit history and monthly repayment statuses reported by other financial institutions. The POS_CASH_balance.csv and credit_card_balance.csv data contain monthly snapshots of the point-of-sale transaction balances, cash loans, and credit card utilization. Additionally, the prior_application.csv table summarizes the applicant's earlier loan applications to the institution, while installments_payments.csv provides detailed records of repayments along with instances of missed repayments on those loans. By bringing together data from these tables, we produce an extensive blend of borrower attributes, such as demographic variables such as age, gender, and housing situation; financial variables such as aggregate income, credit-to-income ratio, and annuity sums; employment characteristics such as days employed and contract form; and behavioral variables such as delinquency counts and punctuality in installment payments. Our dynamic threshold model optimizes profit while ensuring fairness, enhancing risk management and increasing fair access to credit—especially for rural and underbanked populations.

**Use Case Status:** This use case is part of a larger research project.

**Partners:**

Commercial Bank, Credit Bureau, Academic Research Lab

## 2.2 Benefits of the use case

Our AI-driven financial model offers substantial benefits by making credit levels responsive to local economic conditions, such as agricultural production and seasonal income fluctuations. This approach opens up small loans and microfinance to rural families, enabling them to smooth consumption, invest in small enterprises or farm inputs, and accumulate assets. By

doing so, it directly promotes access to basic services and economic resources, helping to reduce rural poverty traps.

The model uses AI-driven, profit-sensitive benchmarks to streamline credit-scoring decisions, driving sustainable economic progress by efficiently channeling capital to qualifying applicants. Banks can capitalize on lower default rates and higher net profits, creating a virtuous cycle of increased lending capacity. This supports development-oriented policies that encourage productive activities and entrepreneurship, while also strengthening the capacity of domestic financial institutions to expand access to banking services.

Our proposal leverages cutting-edge algorithms like Markov Chain Monte Carlo and Bayesian Optimization to design a sound, adaptive financial architecture. By integrating these state-of-the-art AI tools into the core of banking operations, we stimulate risk management innovation and demonstrate how digital technology can enhance service delivery. This approach improves scientific research and promotes innovation, while also expanding access to information and communications technology, working towards universal and affordable access to basic financial services.

Furthermore, our model addresses the issue of static, urban-biased lending by adapting to localized, diverse data, such as community loan schemes, microloan trends, and demographic indicators. This dynamic approach rebalances decision cutoffs to reduce bias, empowering and promoting the social, economic, and political inclusion of all, particularly poor rural borrowers. It also aims to bring down migrant remittance transaction costs and close remittance corridors with high costs, fostering a more equitable financial landscape.

## 2.3 Future Work

Ongoing research will explore widening the scope of the model for application across rural credit networks so that wider-ranging financial decisions become more inclusive. This means including rural-tailored and place-based data sources like field output records, cyclical earnings trends, shadow lending practices (such as collective savings or community loans between individuals), utility payments history, mobile phone money transaction, and outside-bank credit records. Demographic and socio-economic information—such as household makeup, income stability, and dependency on government assistance—will be applied to form more contextually aware borrower profiles. These variables will assist in forming a richer risk assessment environment for groups that have few formal credit histories.

To facilitate deployment in low-data environments, we plan to incorporate semi-supervised learning methods that can operate with minimal labeled data—prevalent in rural areas. The dynamic threshold adjustment logic will also be modified to react to local economic cycles such as crop price fluctuations or times of seasonal unemployment. For situations of small loans or microfinance, thresholds will be made specific to the repayment habits and risk profiles of micro-entrepreneurs, seasonal laborers, and community-based borrowers. Non-conventional metrics such as rent payment histories or remittance flows will be examined for inclusion in the model as other credit indicators.

This growth will include active cooperation with local banks, microfinance institutions (MFIs), cooperatives, and Non-Governmental Organization(NGOs). These collaborators will not just offer access to rural datasets but also insights that make the model representative of on-the-ground realities. Technical work will also comprise creating scalable deployment options using

Docker and Kubernetes, with offline inference support in places with poor internet connectivity. Representational State Transfer Application Programming Interfaces(RESTful APIs) and mobile-friendly lightweight interfaces will be created to enable seamless integration into small-lender setups.

These will include such metrics as enhanced access to credit, better repayment rates, and borrower livelihood shifts (e.g., farm output or business growth). Threshold logic will be regularly updated with real-time economic feedback loops so that the model remains adaptive and equitable.

Lastly, as part of our research objectives, we plan to publish subsequent research highlighting the rural implementation of our model and put it out for peer review. An open-source tool kit consisting of datasets, threshold optimization examples, and region-specific configuration scripts will be published to facilitate wider adoption and community innovation in this area.

## 3 Use Case Requirements

- REQ-01: Access to Loan History Data

Secure access to 5–10 years' anonymized loan application and payment data. They should have numeric fields (e.g., income, amount borrowed, number of days past due), categorical attributes (e.g., job type, education level, type of housing), and time-series (e.g., monthly repayment habits, credit balances over time) to enable comprehensive borrower attribute engineering.

- REQ-02: Data Preprocessing Pipeline

Automated Extract, Transform, Load(ETL) pipelines for filling missing values, normalization/scaling, categorical encoding, and creating derived features (e.g., debt-to-income ratio).

- REQ-03: Computational Infrastructure

Elastic CPU/GPU resources (cloud VMs or on-prem servers) capable of supporting LightGBM training, MCMC sampling (≥1,000 iterations), and Hyperopt searches in a reasonable timeframe.

- REQ-04: MCMC and Bayesian Optimization frameworks

Installed packages for Markov Chain Monte Carlo (such as Numerical Python(NumPy) + Scientific Python(SciPy) or Python Monte Carlo (Version 3)(PyMC3)) and Hyperopt (or equivalent Bayesian optimizer), having same APIs for auto-tuning hyperparameters.

- REQ-05: Integration API

A REST endpoint or batch interface that accepts data from fresh applicants, uses the trained model and dynamic threshold reasoning, and returns immediate "accept/reject" decisions.

- REQ-06: Monitoring & Logging

Ongoing performance monitoring (ROC-AUC, profit scores), input distribution drift detection, and threshold changes noted for auditing and retraining.

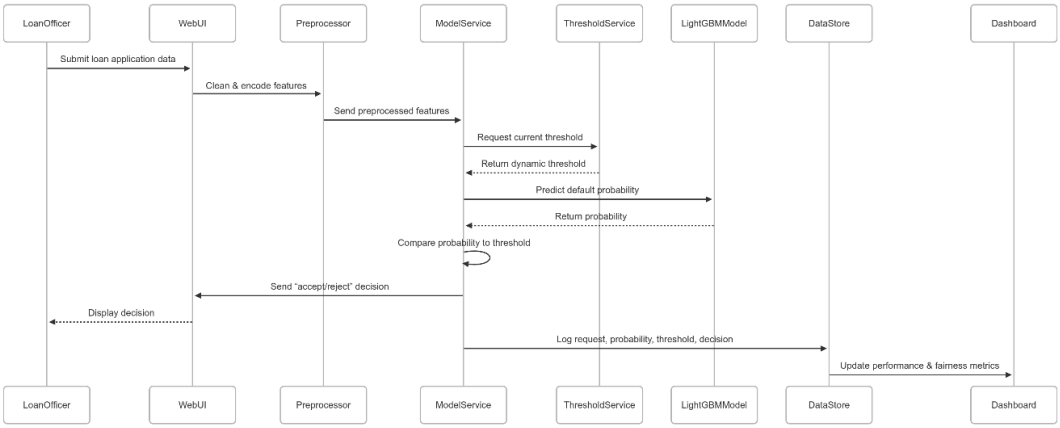
- REQ-07: Compliance and Fairness Testing

Integrated statistical tests (e.g., demographic parity, disparate impact ratio) to ensure threshold updates do not introduce or exacerbate bias, with regular reporting to regulators.

- REQ-08: Visualization Dashboard

Interactive dashboard (e.g., within Google Colab or web User Interface(UI)) displaying current values of threshold levels, profit curves, feature importances, and MCMC trace plots for stakeholder validation.

## 4 Sequence Diagram




## 5 References

[1] https://www.kaggle.com/c/home-credit-default-risk/data: Home Credit Default Risk

[2] https://archive.ics.uci.edu/dataset/144/statlog+german+credit+data: This dataset classifies people described by a set of attributes as good or bad credit risks. Comes in two formats (one all numeric). Also comes with a cost matrix

[3] M. Herasymovych and K. Märka, "Optimizing acceptance threshold in credit scoring using reinforcement learning," *Master's thesis*, Univ. of Tartu, Faculty of Social Sciences, School of Economics and Business Administration, supervised by O. Lukason, 2018.

[4] M. Khashei and A. Mirahmadi, "A soft intelligent risk evaluation model for credit scoring classification," *Int. J. Financ. Stud.*, vol. 3, pp. 411-422, 2015.

[5] J. L. Leevy, J. M. Johnson, J. Hancock, and N. Tran, "Threshold optimization and random undersampling for imbalanced credit card data," J. Big Data, vol. 10, no. 58, pp. 1-13, 2023. doi: 10.1186/s40537-023-00738-z.

[6] E. S. Kamimura, A. R. F. Pinto, and M. S. Nagano, "A recent review on optimisation methods applied to credit scoring models," *Journal of Economics, Finance and Administrative Science*, vol. 28, no. 56, pp. 352-371, 2023. doi: 10.1108/JEFAS-09-2021-0193.

[7] S. Kyeong and J. Shin, "Two-stage credit scoring using Bayesian approach," *J. Big Data*, vol. 9, no. 106, pp. 1-18, 2022. doi: 10.1186/s40537-022-00665-5.

# Use Case 11: InsurVision - Intelligent Insurance Claim Analysis System

**Organization:** Indian Institute of Technology (IIT) Hyderabad

**Country:** India

**Contact Person(s)**:

Shreesh Gupta, cs23mtech12009@iith.ac.in

Arjit Gupta, cs23mtech12001@iith.ac.in

## 1 Use Case Summary Table

| Item | Details |
|---|---|
| Category | Finance |
| Problem Addressed | Manual insurance claim processing is inefficient, with delays that takes days and months, and limited transparency in settlements. |
| Key Aspects of Solution | An Artificial-Intelligence(AI)-driven Agentic system that performs multi-modal analysis (video, image, and document data) to automate claim evaluation and decision workflows. |
| Technology Keywords | Artificial Intelligence, Agents, Large Language Models (LLMs) (GPT 4.1, Gemini 2.5 pro, gemma3), Deep Learning, Natural Language Processing, Automated Workflows. |
| Data Availability | private |
| Metadata (Type of Data) | Text, video, image |
| Model Training and Fine-Tuning | Fine-tuning on a large corpus of historical insurance claims. |
| Testbeds or Pilot Deployments | Our Server(Intel(R) Xeon(R) Platinum 8480+, NVIDIA RTX A6000, RAM 500GB) |

## 2 Use Case Description

### 2.1 Description

InsurVision is a proposed intelligent insurance claiming system designed to revolutionize the conventional claim processing workflow. This solution leverages cutting-edge AI technologies to automate the evaluation of multimedia evidence–videos, images, and documents–submitted by policyholders following an incident.

It should be noted that InsurVision provides 24/7 availability and parallel processing of multiple claims at the same time along with reduced latency.

Also, since we are using LLM's so removing biasness totally is yet an open research problem. So, the results might be biased but human experts can give the feedback that will be used as Feedback mechanism to reduce bias overtime.

Comparison with State Of The Art(SOTA):

There are companies out there(Tractable AI and Lemonade) that use agentic AI services for insurance claims, but by incorporating our Future Work–global rollout with multilingual support and a built-in voice assistant, distributed-ledger data integrity, pilot partnerships alongside open-source collaboration, and specialized deepfake-detection agents–we'll gain an edge.

The system uses:

- Policy Validation Agent: Checking whether claim is valid or not.
- Claim Evaluation Agents: Check all the claim details against the policy, does image analysis, video analysis, exclusion criteria analysis and cost estimation.
- Decision Making Agent: Decides whether the claim is approved or not.
- Output Generation: Generates the Final report.

**Use Case Status:** The implementation plan is in progress.

## 2.2 Benefits of the use case

InsurVision offers significant benefits by integrating advanced AI and deep learning methods into claim processing, thereby modernizing traditional systems and creating a robust digital infrastructure. This innovation enhances operational efficiency and supports the transformation of legacy processes within the insurance sector. The proposal also promotes collaboration among insurance providers, technology developers, and regulatory bodies. By building an open, integrative platform and engaging with global stakeholders, InsurVision aims to create a cooperative ecosystem that addresses common challenges in disaster recovery and risk management. Additionally, accelerated and more transparent claim processing contributes to faster recovery after disasters, ensuring that communities receive prompt support. This efficiency bolsters urban resilience and helps build safer, more sustainable cities in the face of natural and man-made challenges.

## 2.3 Future Work

To enhance our global deployment and localization efforts, we plan to expand the system to support multiple languages, ensuring global accessibility and inclusivity. Additionally, we will introduce a voice assistant agent to make the claiming process more accessible and user-friendly for all users. To ensure data integrity, we will explore the use of distributed ledger technologies to secure and maintain the immutability of claim records, providing a transparent and tamper-proof system.

In building collaborative partnerships, we will establish pilot programs with international insurance firms and disaster response agencies to test and refine our solutions in real-world

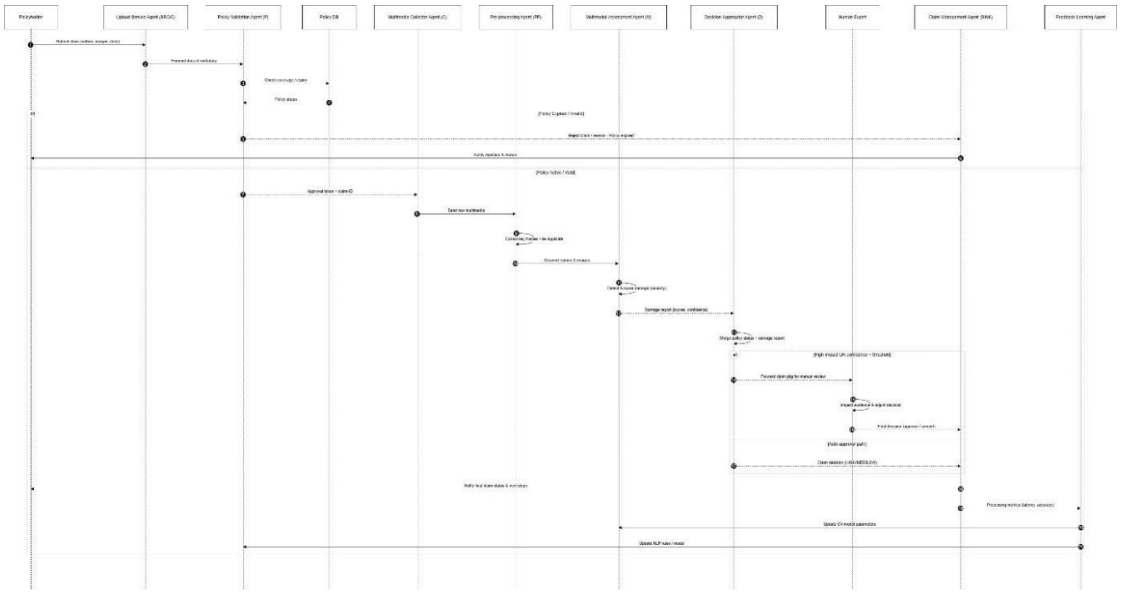
scenarios. We also plan to release key components under an open-source license, encouraging community contributions and fostering further innovation through open collaboration.

To strengthen our system's resilience against fraud, we will incorporate specialized agents capable of detecting manipulated media and deepfakes. This will enhance our ability to identify and prevent fraudulent activities, ensuring the integrity and reliability of our services.

## 3 Use Case Requirements

- **REQ-01**: The system should be capable of processing high-definition video data and documents uploaded by the user.
- **REQ-02**: It must utilize LLM models for comprehensive document analysis and policy validation.
- **REQ-03**: Ensure rapid processing speeds with minimal latency for both image and text analysis components.
- **REQ-04**: The system should be designed to integrate with established risk management standards.
- **REQ-05**: Maintain high system availability and resilience through scalable, cloud-based infrastructure solutions.

## 4 Sequence Diagram



## 5 References

[1] IBM, "AI Agents: Orchestrating Enterprise Workflows," IBM White Paper, 2025. [Online]. Available: https://www.ibm.com/ai-agents

[2] Akira.ai, "Claims Adjustment with AI Agents," Akira AI Blog, Jan. 2025. [Online]. Available: https://www.akira.ai/blog/claims-adjustment-with-ai-agents

[3] Insurance Regulatory and Development Authority of India, "Guidelines on Standard Operating Procedure for Handling Insurance Claims, GEN1189," IRDAI, Tech. Rep., Nov. 2022.

[4] Lemonade Inc., "Claims: How We Do It," [Online]. Available: https://www.lemonade.com/claims

[5] Repairer Driven News, "Tractable AI in Claims Processing: What It Does – and Doesn't Do," Mar. 27, 2024. [Online]. Available: https://www.repairerdrivennews.com/2024/03/27/tractable-ai-in-claims-processing-what-it-does-and-doesnt-do/

[6] https://iith-my.sharepoint.com/:b:/g/personal/cs23mtech12009_iith_ac_in/Ef3NSlM576FNlwJ-Q4eUfzEB2tyUoLgs00cnXITRDUVthw?e=OXai37: Link of Sequence Diagram

# Use Case 12: AI to lift people out of poverty

**Country**: USA (and all countries w/ poverty population)

**Organization**: Era-X Corporation, dba SERA

**Contact Person**(s):

Primary contact: Phong Nguyen, phong@getsera.io

Secondary contact: **Marlon Monroy**, marlon@getsera.io

**Any other contacts:** Mohan Navaratna**,** Mohan@ninegemsinvestments.com

## 1 Use Case Summary Table

| Item | Details |
|---|---|
| Category | Finance |
| Problem Addressed | People in poverty often rely on credit cards to survive. But with low income and high interest rates, escaping that debt is nearly impossible. Worse, many lack the tools or knowledge to manage money effectively, making poor financial health both a root cause and a trap of poverty [1]. |
| Key Aspects of Solution | SERA aims to lift people out of poverty with a double -sided solution: 1- Help pay down CC debt faster without any loan nor extra payment and 2- Enable people to manage and control their finances to build financial stability. We use the Offset Principle to speed up CC debt elimination without extra payment nor good credit, we call it Debt Offset. Furthermore, it uses AI heavily toward their custom built budget so they can manage/improve their spendings through these AI measures: Trimming, Budget ReBalancing, Income Boosting, and Credit Leveraging. Debt Offset solves the immediate CC debt trap, and Smart Budget solves the root cause of debt/poverty: managing finances. |
| Technology Keywords | AI, Agentic AI, LLM. |
| Data Availability | SERA uses users' private data when they connect their checking account and credit card accounts to us. We extract their expenditure data from these accounts to create a custom budget for them. Then we use Debt Offset to help reduce their CC debt faster and Smart Budget measures to help them manage and improve their spending for future stability. Smart Budget uses public data for Trimming, Income Boosting, and Credit Leveraging [6]. For example, SERA uses public yet localized data to recommend using Mint Mobile to reduce their cell phone bills instead of using Verizon or AT&T. It can recommend ways to boost their income in case of needs by doing committed gigs (such as driving for DoorDash or Uber) or uncommitted gigs (such as donating plasma, doing secret shopper), or getting emergency money help from local agencies or social services. With localized data training and continuous user-initiated improvement, it can help anyone in any country [6]. |
| Metadata (Type of Data) | All data is electronic financial information about the users (income, expenditures, credit score, assets and obligations) and their local public financial offerings pertaining to their own situation/qualifications. |

(continued)

| Item | Details |
|---|---|
| Model Training and Fine-Tuning | Model Training & Fine Tuning are very important sections, so we opt for a full description below:<br>SERA's AI engine is designed to optimize financial outcomes for users by supporting debt reduction, spending management, and long-term financial stability. Our model training and fine-tuning process includes the following key stages:<br>1. Data Collection & Preprocessing<br>Source Data: Aggregated, anonymized user financial data including spending patterns, income flows, debt levels, credit card activity, and budget adherence.<br>External Data: Public economic indicators (e.g., inflation rates, consumer spending trends) and behavioral finance research.<br>Privacy & Compliance: All data handling complies with strict standards like GDPR, CCPA, and U.S. financial privacy laws (GLBA, FCRA), ensuring user anonymity and security. It will be localized for any specific country deployment.<br>2. Model Pretraining<br>Initial models are pretrained using general financial behavior datasets based on known public financial offerings as well as our financial tools (Debt Offset, Smart Budget)<br>Focus areas include identifying risk patterns, spending leaks, budget compliance behavior, and debt repayment trends.<br>3. Supervised Fine-Tuning on SERA-Specific Objectives<br>We fine-tune the models using SERA's proprietary dataset collected from early users and simulated financial scenarios.<br>Custom loss functions prioritize outcomes like debt offset optimization, budget discipline enforcement, and predictive cash flow management.<br>4. Reinforcement Learning with Human Feedback (RLHF)<br>Peer users including special users such as Financial coaches, economists, and behavioral scientists manually review and guide model outputs to improve advice relevance, user personalization, and decision-making support.<br>This feedback loop ensures the AI continually learns to provide practical, actionable, and empathetic recommendations.<br>5. Continuous Personalization<br>Once deployed, each user's interaction further trains the AI engine in real-time through adaptive learning.<br>The system dynamically adjusts its guidance based on individual user progress, setbacks, lifestyle changes, and financial emergencies.<br>6. Bias Mitigation and Risk Monitoring<br>Regular audits are conducted to monitor for model bias, especially to prevent systemic disadvantages based on income, background, or geography.<br>Robust fallback strategies are designed to prevent poor advice during periods of economic volatility.<br>7. Model Evolution for Future Expansion<br>Beyond Debt Offset, SERA's AI is being expanded to address savings growth, investment starter strategies, and microbusiness cash flow management, creating an end-to-end AI-powered financial wellness ecosystem. |

# 2 Use Case Description

## 2.1 Description

There are 165M Americans who carry credit card debt for over a year, and 81M of them live near or below the poverty line. Consumer debt is the main catalyst keeping people in poverty. Moreover, regardless of how people get into poverty, what keeps people in poverty in each country is equally important. There is no current tool to help people manage and improve their spending patterns, leading to their current poverty situation [6].

SERA aims to eliminate credit card debt faster without any loans or extra payments while helping to manage/improve people's spending habits for future stability. Currently created for North and South America, it can be localized to any country that can use banking products like checking, savings accounts, credit/debit cards, etc. [7].

[Describe the innovative technological approach being implemented.] SERA uses the Offset Principle, a powerful yet little-known financial concept in its core product, Debt Offset to help reduce CC debt faster than any current debt service products today. Moreover, SERA also addresses the root cause of debt by managing and improving people's spending patterns through the use of AI to enforce a balanced budget.

[Explain the expected impact and benefits of the solution.] With the current debt service model, there are only 2 legitimate ways to pay down your debt faster. One is to pay more than the monthly minimum (which requires more disposable income) and two is to lower the interest rate of your debt (which requires good credit). For the first time, working people can reduce their CC debt faster without the need for more disposable income nor good credit score. In particular, this new solution will benefit those who have low income and poor credit the most, as they have no other viable debt solution today (81M Americans alone). Once CC debt is paid off, the funds from previous CC payments can be used toward reducing any other deb,t such as Student Loan, Consumer Loan, Car Loan, etc. Moreover, the SERA work in helping people improve their spending patterns can also be used by any working adult to improve their finances (with or without credit card debt) to gradually build their financial stability (worldwide impact as any working adult anywhere can manage and improve their economic status with our Smart Budget tool) [4][5].

## 2.2 Benefits of the use case

- No Poverty. SERA helps lift people out of poverty by first helping them to escape their CC debt faster without extra payments nor good credit (currently not possible). Their biggest problem is debt, especially CC debt as it's easier to get and CC is needed for their daily spending [5]. More importantly, SERA will then help them manage and improve their finances so that they can stay out of debt to gradually build their financial stability. Note that "improving" finances might be in the form of boosting their current income with localized opportunities that they might not be aware of. For example, a US citizen is short of $150 for rent within a few days, what can he do? Our Income Boosting feature can recommend a committed gig (driving Uber or Doordash), or an uncommitted gig such as selling items on eBay, donating plasma or doing a secret shopper. If they are consistently short of funds, they might want to adopt a more permanent income boosting method recommended by SERA.
- Economic Growth. Low-income people do not have a pathway to increase their economic level due to debt and a lack of tools to improve their current financial situation. Many

people remain within their paycheck-to-paycheck mode (and borrowing in between) because there is no method to control and improve their spending patterns. SERA helps not only with debt elimination but also with improvement to their finances so they can have the opportunity to elevate their economic status.

- Reduced Inequalities. As in most countries, the poor get poorer due to limitations of their economic mobility, and of course, the rich can get richer due to access to economic opportunities that others can not touch. This inequality gap is widening in every country due to systematic constraints built into our banking system and their products. SERA provides every working adult with a new set of tools (starting with Debt Offset and Smart Budget) to level the playing field for those who have lower incomes and poor credit. This equitable opportunity to economically advance themselves, first with their current debt and then onward to build their financial stability through spending improvements, transcends across all population segments regardless of race, gender orientation, geographical environment, economic status, or physical limitation. Using the Offset Principle (which is the core of Debt Offset), SERA will fundamentally change the way all banking systems work from their current FLOAT model to the PEOPLE BANKING model to benefit the working class. This is a deep subject that needs more than just a short description, but here is roughly how it works. In Debt Offset, SERA calculates how much money a user would normally use in the month through their credit cards, and we call it the Monthly Credit Transaction Amount, or MCXA. Once they earn their income, we move the MXCA from their checking account into their CC balance, resulting in a lower daily balance and faster reduction without extra payment. This MCXA can still be used in the month for their normal spending from this CC balance, but the longer it stays in their CC account, the bigger the offset (or savings) the user can benefit from it. Fundamentally, this new Offset Method will transform the current bank's float system, where banks use their clients' deposits to benefit themselves, to a new paradigm where the clients' deposits will be used to improve their own financial obligations first.

**Partners**

We are currently in talks with the following entities to maximize SERA's exposure to empower people of all walks of life.

- United for ALICE (unitedforalice.org): We are working to bring our innovative solution to empower the ALICE people to achieve financial stability. Here is their description of United Way's ALICE initiative:

"United For ALICE is a driver of innovation, research, and action to improve life across the country for ALICE (Asset Limited, Income Constrained, Employed) and for all. The development of the ALICE measurements, which provide a comprehensive, unbiased picture of financial hardship, has revealed that 41% of U.S. households cannot make ends meet has emerged.

The ALICE research quantifies and describes the number of households that are struggling financially. From humble beginnings as a study to understand the struggles of families in one New Jersey county, this body of research has worked to increasingly shed light on a hidden population we now know as ALICE. Harnessing this data and research on the mismatch between low-paying jobs and the cost of survival, ALICE partners convene, advocate, and collaborate on solutions that promote financial stability at the local, state, and national levels.

After a decade of work, the spark ignited by the initial study has grown into a grassroots ALICE movement across more than two-thirds of U.S. states to change the national dialogue about financial hardship. Led by United Way of Northern New Jersey, the ALICE movement includes United Ways, corporations, nonprofits, and foundations in Alabama, Arkansas, Colorado, Connecticut, Delaware, District of Columbia, Florida, Georgia, Hawai'i, Idaho, Illinois,

Indiana, Iowa, Kansas, Kentucky, Louisiana, Maine, Maryland, Michigan, Minnesota, Mississippi, New Jersey, New Mexico, New York, North Carolina, Ohio, Oklahoma, Oregon, Pennsylvania, South Carolina, Tennessee, Texas, Virginia, Washington, West Virginia , and Wisconsin; we are United For ALICE."

- StandTogether.Org We aim to partner with StandTogether to bring our solution to as many people as possible to bridge the income gap among the poor and unserved folks and their economic plight through their vast network of constituents. One of their initiatives is to solve the Poverty problem once and for all in the US and beyond. This is the description of StandTogether:

"Stand Together is a philanthropic community that helps America's boldest changemakers tackle the root causes of our country's biggest problems, from education to the economy, broken communities, and toxic division, among dozens of other pressing issues. We provide our partners with access to resources, including funding, a unique network of peers, and a playbook for applying proven principles to transform lives and society. Complex problems require a comprehensive approach. That's why the Stand Together community partners with a wide diversity of changemakers, including thousands of nonprofit leaders, Fortune 50 CEOs and philanthropists, NFL legends, civil rights leaders, Grammy-winning musicians, and grassroots policy activists."

- Operation Hope (Operationhope.Org) We are working to equip Operation Hope with our powerful tool to practically bring financial education to people's daily lives so they can learn with real life action from SERA to deal with debt and spending habits. This is the description of Operation Hope:

"Operation HOPE is a nonprofit organization dedicated to advancing economic opportunity and financial literacy for underserved communities. Through its comprehensive programs, Operation HOPE empowers individuals to achieve financial stability and independence within over 200 locations across the United States (often inside bank branches). They provide financial literacy programs such as Credit & Money Management, Homeownership Program, Youth and Young adult Program, and other financial services and education to local folks."

## 3 Use Case Requirements

- **REQ-01**: It is critical that the system supports secure integrations with banking platforms to access real-time financial data while maintaining end-to-end encryption.
- **REQ-02**: It is critical that a multi-layered LLM architecture is implemented, incorporating classification, behavioral detection, recommendation, and orchestration layers.
- **REQ-03**: It is expected that the platform operates on a cloud-native, encrypted infrastructure that complies with modern security standards.
- **REQ-04**: It is critical that the system supports real-time monitoring and automated financial nudging/enforcement actions based on user behavior.
- **REQ-05**: It is expected that high availability of 99.9% uptime is maintained to ensure uninterrupted access and reliability for users.
- **REQ-06**: It is expected that continuous AI feedback loops are established to enable ongoing learning and improvement of model outputs.
- **REQ-07**: It is critical that the system provides users with clear, explainable communications, especially for budget decisions and financial advice.
- **REQ-08**: It is critical that dynamic user profiling is enabled, allowing real-time updates based on spending patterns and evolving financial goals.

- **REQ-09**: It is expected that the system includes full auditability and adheres to compliance standards such as PCI DSS and SOC 2 to build trust and transparency.
- **REQ-10**: It is of added value that the system delivers dynamic budget rebalancing to align user spending with predefined financial goals.

## 3.1 Future Work

- [Describe planned enhancements, developments, or next steps.] Regarding our AI engine, continuous improvement is needed to train the model with the latest financial product offerings, tools and innovations as well as users' input [6]. However, as suggested, we will fine-tune our engine for achieving further financial goals such as Emergency Fund, Asset Building or Passive Income Accumulation per our financial freedom roadmap. Each step requires SERA to be trained with entirely different datasets regarding its objective. For example, Emergency Fund initiatives can be different for each user as it can be internal (use savings from my paycheck or spending to create my own emergency fund) or external (use a portion of my fund to offset other people's debt in exchange for % of return), or combination of both. Similarly, for Asset Building or Passive Income Accumulation, it needs datasets for investment vehicles, asset types, risk tolerance, time frame constraint, and income type (asset protection/building vs income producing), etc.
- [Outline additional resources needed, such as technology upgrades, further data collection, model improvements, etc.] Here is our detailed write-up for needed resources:

### Model Layer

- Our plan is to utilize PyTorch mainly for all our Machine Learning and AI needs
- Base LLM Models:
- Mixtral or LLaMA for context handling
- Specialized models: TimeLLM, T5, Longformer (for large document/time modeling).
- Behavior Prediction Models:
- Train smaller classifiers (XGBoost, LightGBM) for quick pattern detection.
- Use Hybrid LLM + Classical ML when speed is needed.
- Use La Plateforme for finetuning needs

### Data Layer

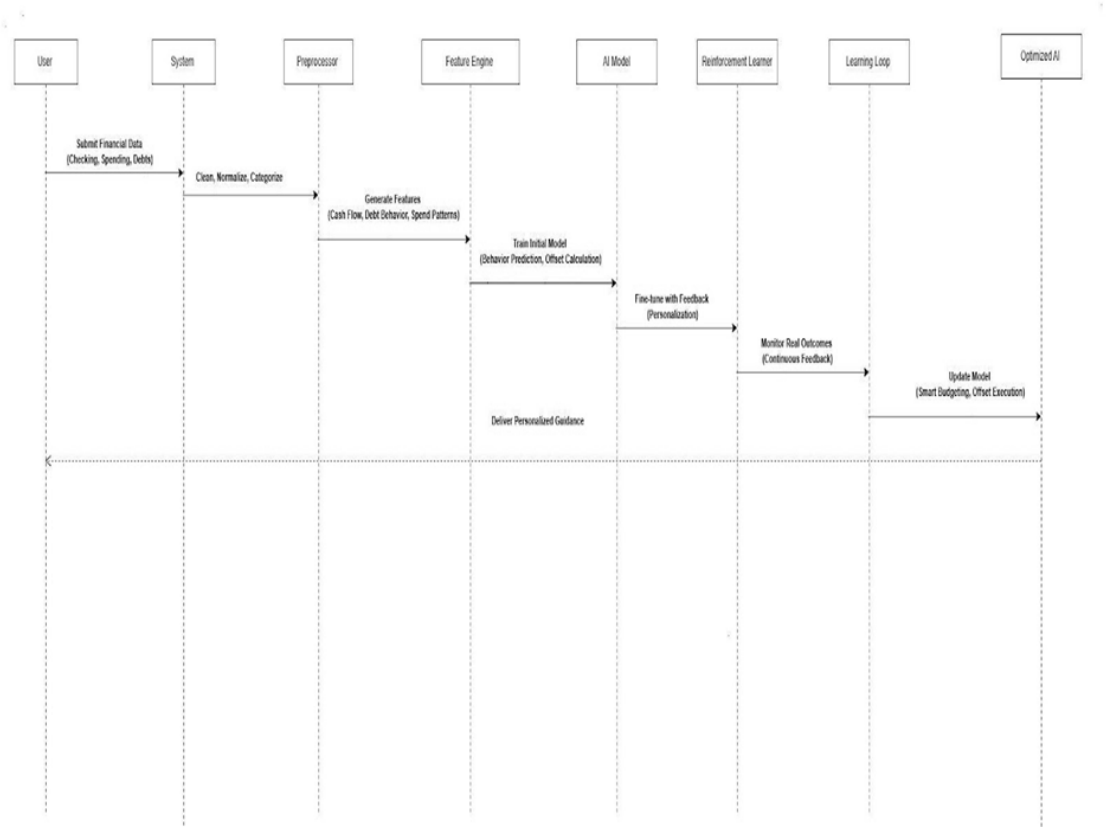
- Sequential Data Pipelines:
- Preprocess transaction streams into chronological sequences.
- Windowing: Break sequences into 1-week, 1-month, 3-month snapshots.
- Behavioral Embedding Store:
- Vector DB for storing user behavioral profiles.
- Compare "current self" vs "historical self" for anomaly scoring.

### Infrastructure Layer

- Model Training Compute:
- Heavier GPU (A100 or H100) for fine-tuning on long-context data.
- Use cloud like AWS Sagemaker, Hugging Face, or local clusters if needed.
- Serving and Detection Engines:

- Real-time ingestion (Kafka, PubSub) for new transaction streams.
- Microservices using FastAPI or Node.js for fast predictions.
- Behavior Monitoring Platform:
- Dashboards (Superset, Metabase) to visualize emerging patterns across users.
- Alerting pipelines (PagerDuty, Slack) if massive pattern shifts occur.
- In addition to our current discussions with large non-profit organizations such as United Way and Financial Health Networks, we are eager to collaborate with international organizations like the Bill & Melinda Gates Foundation and World Vision to help combat poverty and create better economic opportunities worldwide. We also aim to expand our global reach and credibility to serve a broader audience beyond the Americas [7][8].

## 4 Sequence Diagram




## 5 References

[1] Federal Reserve Bank of Philadelphia, "Moving Up or Running in Place? Examining Economic Stability and Mobility in Philadelphia." [Online]. Available: https://www.philadelphiafed.org/community-development/workforce-and-economic-development/moving-up-or-running-in-place-examining-economic-stability-and-mobility-in-philadelphia

[2] Federal Reserve Bank of Philadelphia, "Savings versus Debt: The Effects of Survey Question Order on Consumers' Reported Financial Priorities." [Online]. Available: https://www.philadelphiafed.org/consumer-finance/consumer-credit/savings-versus-debt-the-effects-of-survey-question-order-on-consumers-reported-financial-priorities

[3] Federal Reserve Bank of St. Louis, "Revolving Consumer Credit Owned and Securitized, Outstanding," FRED Economic Data. [Online]. Available: https://fred.stlouisfed.org/series/RCCCBSHRMIN

[4] TransUnion, "2025 Consumer Credit Forecast." [Online]. Available: https://newsroom.transunion.com/2025-Consumer-Credit-Forecast/

[5] TransUnion, "Q1 2025 Credit Industry Insights Report (CIIR)." [Online]. Available: https://newsroom.transunion.com/q1-2025-ciir/

[6] G. Elliehausen, "Written Statement," CFPB Symposium on Behavioral Economics, Consumer Financial Protection Bureau, 2020. [Online]. Available: https://files.consumerfinance.gov/f/documents/cfpb_elliehausen-written-statement_symposium-behavioral-economics.pdf

[7] Testbeds or Pilot Deployments. [Online]. Available: https://getsera.io

[8] Code repositories. [Online]. Available: https://github.com/eraxadr

# 4.7. Category 7: Education

## Use case 1: AI-Powered Math Teaching Assistant - Process-level Feedback & Assessment

**Organization:** ZEZEDU Corp.

**Country:** Republic of Korea

**Contact Person(s)**:

Minji Cha (minji@zezedu.com)

Jujin Lee (jujin@zezedu.com)

### 1 Use Case Summary Table

| Category | Education |
|---|---|
| The problem to be addressed | • Providing frequent, high-quality feedback in math education is essential but highly time-consuming for teachers in traditional education models<br>• Digital platforms primarily analyse students' final answers without considering the problem-solving process, thus offering limited educational benefit |
| Key aspects of the solution | • Specialized AI that comprehends handwritten math equations and evaluates the complete problem-solving process<br>• Immediate, actionable, process-level feedback tailored for both students and educators, reducing educators' manual workload while ensuring fair evaluation<br>• Enhanced access to high-quality, personalized math education, bridging educational divides regardless of geographic location and economic background |
| Technology keywords | AI for Education, Handwriting Equation Recognition / Optical Character Recognition (OCR), Process-level Feedback Generation, Mathematical Problem Solving, Automated Grading and Assessment, Clustering, Personalized Learning |
| Data availability | Private dataset |
| Metadata (type of data) | Data from students' learning history, including both textual and image data |

(continued)

| Category | Education |
|---|---|
| Model training and fine-tuning | • NLP<br>• OCR<br>• Generative AI (Open Source LLMs)<br>• RAG |
| Testbeds or pilot deployments | Can test demo from official website: cherrypot.net [4] |

# 2 Use Case Description

## 2.1 Description

Providing high-quality math education requires frequent, process-level feedback to help students understand not just what they got wrong, but why. However, in public education settings where one teacher is responsible for many students, delivering this level of personalized feedback consistently is extremely challenging. As a result, such detailed support has traditionally been available only through costly private tutoring, making it inaccessible to many.

To address this challenge, we developed proprietary AI models that recognize students' handwritten math solutions and analyze their problem-solving processes. Unlike most digital learning platforms that focus solely on final answers, CherryPot analyzes each student's full problem-solving process to provide detailed, process-level feedback, tailored follow-up questions, and personalized learning paths. For teachers, we support partial score assessment, reducing repetitive grading tasks and allowing them to focus more on instruction. By delivering high-quality, individualized feedback at scale and at a lower cost, we help close educational gaps and ensure equitable access to quality math education, regardless of geographic location.

CherryPot supports students from Grade 3 to Grade 11 and is accessible across a range of devices, including mobile phones, tablets, and PCs. Students can submit their work by taking photos (mobile), writing directly on the device (tablet), or typing (PC). We support a variety of curricula, including Korea's national curriculum, U.S. state-specific standards, and international curricula such as IB and IGCSE.

Our platform is a school-proven solution, trusted by over 190 public schools across Korea. Usage data shows meaningful academic improvement among students after just one month of engagement. To extend our impact globally, we are now conducting pilot programs with schools in the United States.

Use case status:

- Deployed and scaling globally
- Proven success in real-word educational field, used in 190+ public schools in Korea
- 4000+ active users
- Pilot testing underway in U.S. public schools

**Partners**

Seoul National University-Affiliated School Promotion Institute

## 2.2 Benefits of use case

CherryPot makes high-quality math education accessible to all students, regardless of their geographic location or socioeconomic status. By leveraging AI to provide personalized, process-level feedback and automated grading, our platform delivers a level of educational support that was once limited to students with access to expensive private tutors or advanced academic resources. During the COVID-19 era, we also collaborated with various government ministries to launch diverse initiatives, further expanding our impact. This allows for more equitable learning opportunities and helps close the achievement gap.

Our AI grading system ensures consistent, unbiased assessment of student work, helping reduce disparities in educational evaluation. Our AI model provides equal treatment for all students, regardless of background. Additionally, the affordability and scalability of our solution allows under-resourced schools and regions to benefit from the same high-quality feedback as well-funded institutions.

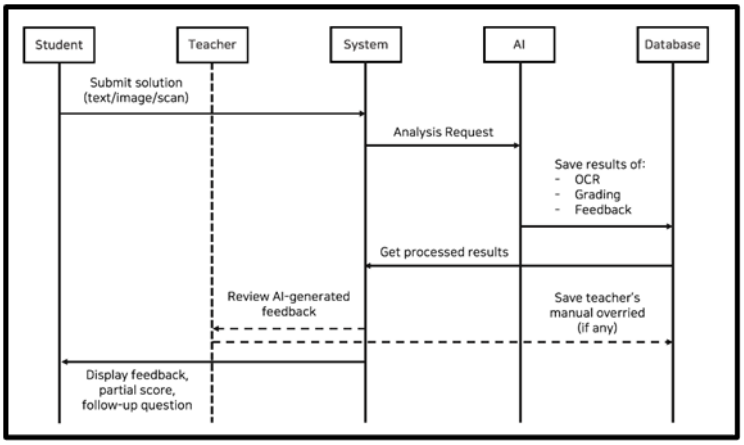
## 2.3 Future Work

- Data Collection and Model Improvement: Continuously gather data to boost accuracy, adaptability, and quality, while harnessing a virtuous cycle where teacher reviews are systematically fed back into model improvements.
- Multilingual Support: Extend the AI model to support multiple languages and mathematical symbols used in different countries, enabling global curriculum coverage.
- Advanced Learning Dashboard: Provide detailed student analytics to help educators track progress, identify learning gaps, and receive AI-generated instructional suggestions.
- Offline Mode Development: Develop an offline-compatible version of the platform to support students in regions with limited internet access, such as parts of Africa and other underserved areas.

# 3 Use Case Requirements

- **REQ-01**: It is required for users to register and complete the authentication process.
- **REQ-02**: The system should support various student inputs, including handwritten image uploads, typed text, and scanned paper submissions.
- **REQ-03**: The system should support real-time processing capabilities in order to deliver immediate feedback to students and teachers.
- **REQ-04**: The system should support educator-side interfaces for reviewing AI grading results and providing manual overrides when necessary.

## 4 Sequence Diagram




## 5 References

[1] Study on the effectiveness of customized mentoring using edutech [link 1] [link 2]

[2] Singaporean EdTech Accelerator EduSpaze Announces Seventh Cohort of Startups [link]

[3] ZEZEDU revolutionises math education with its AI-powered tools, helping students to excel [link]

[4] Test deployment of the platform [link]

## Use case 2: QANDA – An AI Tutor for everyone, advancing equity in education

**Country:** Republic of Korea

**Organization:** Mathpresso Inc.

**Contact Person(s)**:

Monsoo Jung (jason.jung@mathpresso.com)

Donghun Oh (kilian.oh@mathpresso.com)

Mathpresso Support Team (support@mathpresso.com)

### 1 Use Case Summary Table

| Category | Education |
|---|---|
| The problem to be addressed | Access to quality education remains highly unequal across and within countries, worsened by rapid technological advancement. Many students lack teachers to ask questions or receive personalized guidance, limiting their learning opportunities. This inequity obstructs progress towards enhancing quality education, reduced inequalities, and decent work and economic growth. |
| Key aspects of the solution | QANDA provides an AI-powered learning platform that offers real-time question-and-answer support through a combination of optical character recognition (OCR) and large language models (LLMs). Beyond simple answers, the platform delivers a personalized learning experience by analyzing student behavior, curriculum data, and study patterns. QANDA's AI agents support not only Q&A but also note-taking, memorization, self-evaluation, and concept mastery to enhance self-directed learning. |
| Technology keywords | Generative AI (GenAI), Large Language Models (LLMs), Optical Character Recognition (OCR), Retrieval-Augmented Generation (RAG) |
| Data availability | Private dataset |
| Metadata (type of data) | Text, Image |
| Model training and fine-tuning | QANDA utilizes large language models (LLMs) fine-tuned for educational tasks, enhanced with a retrieval-augmented generation (RAG) pipeline to ensure factual accuracy. The platform combines optical character recognition (OCR) for image-based input processing and zero-shot reasoning for handling unseen questions. AI models continuously improve through learning data accumulation, enabling personalized guidance and scalable content generation across diverse subjects. |

# 2 Use Case Description

## 2.1 Description

Educational inequity remains a major challenge globally, worsened by limited teacher access and uneven technological adoption. This threatens progress on enhancing quality education, reduced inequalities, and decent work and economic growth. QANDA aims to democratize quality education by offering an affordable, AI-driven personal tutor, improving access, learning efficiency, and motivation for all students. Combining OCR, large language models (LLMs), and a RAG pipeline, QANDA provides real-time question-solving, adaptive feedback, and personalized study plans through a mobile-first platform. QANDA empowers students with self-directed learning tools, narrowing educational gaps and enabling broader economic and social inclusion.

**Partners**

- Technology and Research Partners: Technology partners and investors (Samsung, Google, Bytedance, Meta) and research institutes (Yonsei University, Arizona State University) support AI infrastructure, model development, and educational impact research.
- Government and Content Partners: Collaborations with Korean government ministries and leading publishers drive public school deployments and localized content development.

## 2.2 Benefits of use case

QANDA provides affordable, AI-powered personalized tutoring, ensuring inclusive and equitable education opportunities for students regardless of their background.

The platform addresses educational disparities both across countries and within societies, particularly benefiting underserved and low-income communities

By enhancing access to education, especially in emerging economies, QANDA helps more students acquire the skills necessary for decent work and sustainable economic growth.

## 2.3 Future Work

- Expansion of AI Tutor Capabilities Across Learning Behaviours

QANDA is expanding its AI Tutor's role beyond question-and-answer interactions to cover the full learning journey. The AI Tutor will actively assist students in notetaking, summarizing study materials, memorizing key concepts, conducting self-evaluation, and reinforcing mastery through continuous personalized feedback and guidance.

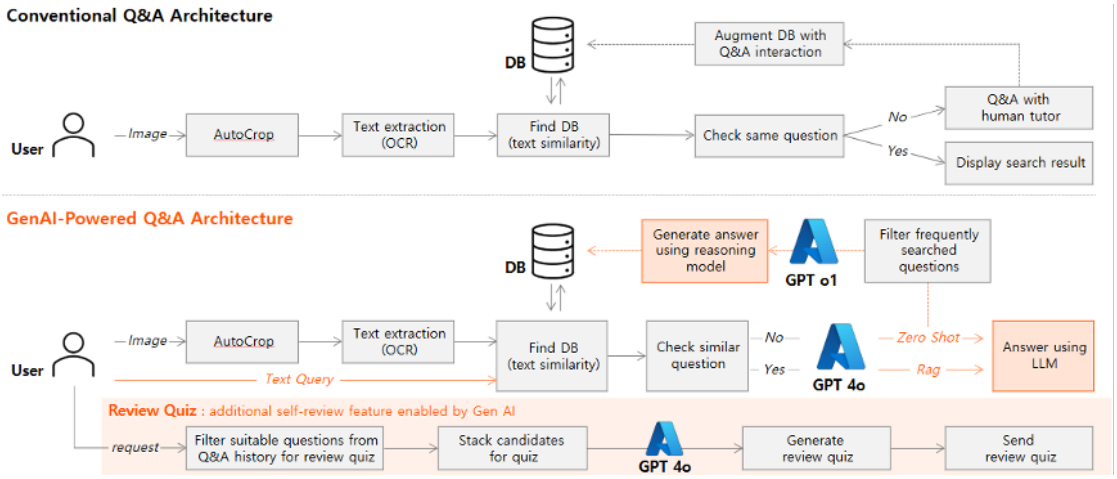
- Geographic Expansion with Hyperlocal Content Strategy

As education systems differ greatly across regions, QANDA is strengthening its localization efforts by implementing a hyperlocal content acquisition strategy. Content will be adapted not only through partnerships with local educational institutions and publishers, but also dynamically sourced from student and teacher feedback behaviors within the app. This ensures that QANDA's AI Tutor delivers culturally relevant, curriculum-aligned support optimized for each market's unique educational needs.

## 3 Use Case Requirements

- **REQ-01**: It is critical to implement a high-accuracy Optical Character Recognition (OCR) system for accurate extraction of text from images and handwritten notes.
- **REQ-02**: It is critical to integrate large-scale Large Language Models (LLMs) and design an effective user interface to deliver precise, contextual, and accessible educational support.
- **REQ-03**: It is critical to utilize a Retrieval-Augmented Generation (RAG) system to ensure AI responses are grounded in verified educational databases.
- **REQ-04**: It is critical to deploy an adaptive learning analytics engine to personalize learning journeys based on student behavior and study patterns.
- **REQ-05**: It is critical to maintain a localization framework to adapt content effectively across different languages, curricula, and cultural contexts.

## 4 Sequence Diagram




## 5 Reference

[1] Internal references

# Use case 3: AI Helpy for AI Textbook

**Country:** South Korea

**Organization:** Elice, Inc.

**Contact Person(s)**:

Suin Kim (suin@elicer.com)

Vu-Linh Le (drew.le@elicer.com)

## 1 Use Case Summary Table

| Category | Education |
|---|---|
| The problem to be addressed | • Using generative AI chatbots in digital textbooks for Korean public educational institutions.<br>• It is imperative that schools in Korea provide a safe learning environment within the school curriculum. |
| Key aspects of the solution | • AI models specialized for education.<br>• AI tutors have been shown to enhance the educational experience by personalising lessons for students and reducing teachers' workloads. |
| Technology keywords | Custom LLM, RAG |
| Data availability | Private dataset |
| Metadata (type of data) | Text only, documents with images are supported through OCR |
| Model training and fine-tuning | Open-source LLMs are fine-tuned with an extensive amount of Korean text data in various domains: Conversational, instruction following, safety guide-line, and education-specific. |
| Testbeds or pilot deployments | [2] |

## 2 Use Case Description

### 2.1 Description

AI Helpy is a generative AI tutor developed by Elice. Elice, Inc. has created AI textbooks for Korean school education, integrating AI Helpy to enhance these materials. 'AI Helpy for Textbook' connects small AI models in a pipeline to create a safe learning environment for students. Previously, teachers using traditional textbooks found it challenging to provide

personalized 1-on-1 feedback. With AI Helpy, teachers can reduce their workload while offering students more detailed, tailored feedback.

The system incorporates a Toxicity Checker model to block inappropriate language, ensuring students can safely learn and understand their role as responsible global citizens. It also includes a model to detect hallucinations, which helps reduce misinformation, ensuring the delivery of accurate information. The Pre-learning Detection model prevents students from engaging in content beyond the textbook, encouraging them to follow the curriculum naturally. This ensures all students start from an equal point, promoting fair and balanced educational opportunities. By limiting pre-learning, students can deepen their understanding of current material, reducing stress and fostering a healthier learning experience.

Additionally, the Tone Changer model adapts responses to students' age groups, offering kind and appropriate feedback. It also provides personalized guidance on spelling errors, helping students stay engaged and follow lessons effectively.

Integrating AI into school education is a significant innovation that enhances education quality and supports a sustainable system. By offering customized feedback based on students' individual levels and speeds, it provides equal educational opportunities for more students, improving both the efficiency and quality of learning.

AI Helpy officially supports Korean, English, Japanese, Thai, and Vietnamese. This is an open knowledge base possible. This system is provided two-way (Student Mode, Teacher Mode), with expert supervision on the teacher side when AI Helpy interacts with students.

## 2.2 Benefits of use case

AI Helpy, applied to the AI-based platform "AI Textbook," provides high-quality education as an equal opportunity for all. By offering content that reflects language translation and cultural differences through AI Helpy technology, it ensures that a wider range of students can receive equitable education.

By reducing teachers' workloads and improving their working environment, it allows them to focus on the essence of education

Furthermore, it promotes the digitization and efficiency of education, positioning technology as a tool for sustainable development.

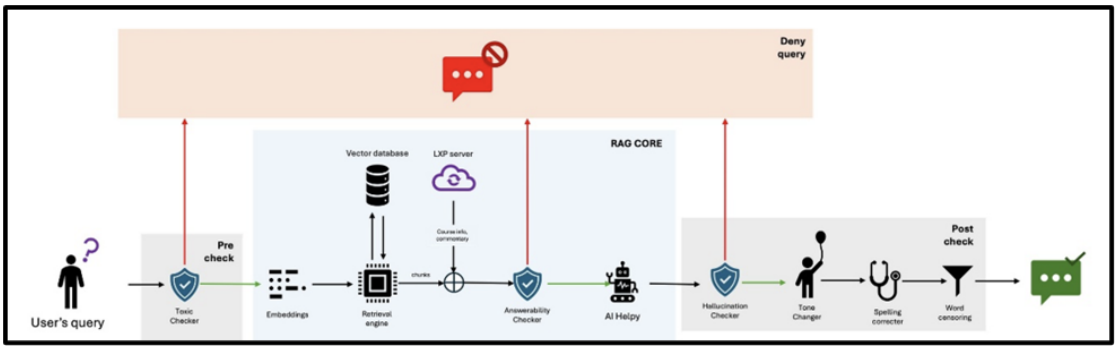
## 2.3 Future Work

To further advance the AI Helpy for textbooks, we plan to diversify its learning by incorporating data from various countries, enabling the platform to expand internationally. Additionally, we aim to enhance subject-specific data learning to improve the quality of personalized 1-on-1 feedback, allowing students to achieve better learning outcomes through more interactive engagement with AI Helpy. By integrating additional small AI models and optimizing the operational pipeline for faster performance, we will deploy the system on a secure cloud platform to ensure scalability and reliability. This model development will actively support students' engagement with school education, helping them achieve measurable academic success.

## 3 Use Case Requirements

- **REQ-01**: For text generation, we would require the following resources:
  GPU: Nvidia A100 (80 GB vRAM PCIe) x 2
  CPU: 64vCore, Memory: 64 GiB
- **REQ-02**: For the toxic check model, we would require the following resources:
  GPU: A100 80GB PCIe MIG 1g-10GB
  CPU: 2vCore, Memory: 24 GiB
- **REQ-03**: For the tone change model, we would require the following resources:
  GPU: A100 80GB PCIe MIG 1g-10GB
  CPU: 2vCore, Memory: 24 GiB
- **REQ-04**: For the relevance check model, we would require the following resources:
  GPU: A100 80GB PCIe MIG 3g-40GB
  CPU: 2vCore, Memory: 24 GiB
- **REQ-05**: For the vector store, we would require the following resources:
  Milvus vector store + Google Cloud Storage
  vCU: Auto-scaling vCU and Storage

## 4 Sequence Diagram




## 5 References

[1] https://elice.io/en/newsroom/eliceschool-aidt

[2] https://elice-library.elice.io - Pilot Deployment

# Use case -4: AI-Powered Offline Learning for Higher Education in Crisis Zones

**Country:** Palestine

**Organization:** Al-Aqsa University, Gaza

**Contact Person(s)**:

Shadia Y. M. Baroud (sy.baroud@alaqsa.edu.ps)

Abdelrafe Elzamly (abd_elzamly@alaqsa.edu.ps) - Head of Department of AI and Computer Science

## 1 Use Case Summary Table

| Category | Education |
|---|---|
| The problem to be addressed | War-induced infrastructure damage disrupts higher education; students lack reliable power, internet and personal computers. |
| Key aspects of the solution | Education (Education technology, Accessibility, Resilient learning, Offline AI)<br>• Agent-based AI (Multi-agent Systems), Wi Fi mesh, knowledge base<br>• Portable Raspberry Pi "AI servers" host content & tutors offline.<br>• Mobile-first access via local Wi-Fi<br>• AI-driven content recommendation & adaptive tutoring.<br>• On-device auto-grading and analytics. |
| Technology keywords | AI agents, Offline AI, Edge computing, NLP tutor, Low-power hardware, Adaptive learning, Raspberry Pi, Solar-powered servers. |
| Data availability | **Mixed:** Public OER datasets; Local student-interaction data (private, stored on-device). |
| Metadata (type of data) | Text, code snippets, speech (TTS/STT), images/diagrams, assessment logs. |
| Model training and fine-tuning | Lightweight transformer-based NLP models (e.g., DistilBERT) fine-tuned offline; Matrix-factorisation for recommendations; on-device quantization (TensorFlow Lite). |
| Code Repositories | University Moodle, Internal GitLab (private); plan to open-source core modules after security review. |

# 2 Use Case Description

## 2.1 Description

In Gaza, ongoing conflict and infrastructure destruction have led to frequent electricity blackouts and almost nonexistent internet connectivity. More than 90,000 university students have lost access to conventional e-learning platforms that require stable connections and personal computers. The objective of this use case is to deliver an "AI-powered, fully offline learning environment" that runs on low-power hardware and is accessible from any smartphone via local Wi-Fi. The innovative approach combines lightweight NLP tutors, AI-driven content recommendation and automated assessment – all executed on-device – to ensure uninterrupted, personalized learning during crises.

The solution now uses autonomous micro-agents that communicate over a Wi-Fi mesh. A knowledge-base agent manages educational content, allowing periodic updates independent of the AI framework [1]. If any server receives fresh material, it shares it across the mesh so every hub stays current. English remains the initial language; Phase 2 will integrate Arabic GPT models for full bilingual support. The design follows the decentralised multi agent coordination pattern demonstrated in [2], [3], [4].

Expected benefits include 24/7 learning continuity, reduced instructor workload through auto-grading and scalable deployment to other conflict or disaster zones.

**Use Case Status:** Seeking partners for scale-up.

### Partners

- Al-Aqsa University – Lead institution, deployment & evaluation
- ITU
- Local NGOs & Educational Foundations – Hardware sponsorship, community training
- Open-education content providers – Curated OER materials
- Potential tech partners – Solar and low-power hardware suppliers
- Potential tech partners – Solar and low-power hardware suppliers

## 2.2 Benefits of use case

- Maintains higher-education access despite infrastructure collapse.
- Demonstrates innovative edge-AI in low-resource settings.
- Level the playing field for students lacking devices or connectivity.

## 2.3 Future Work

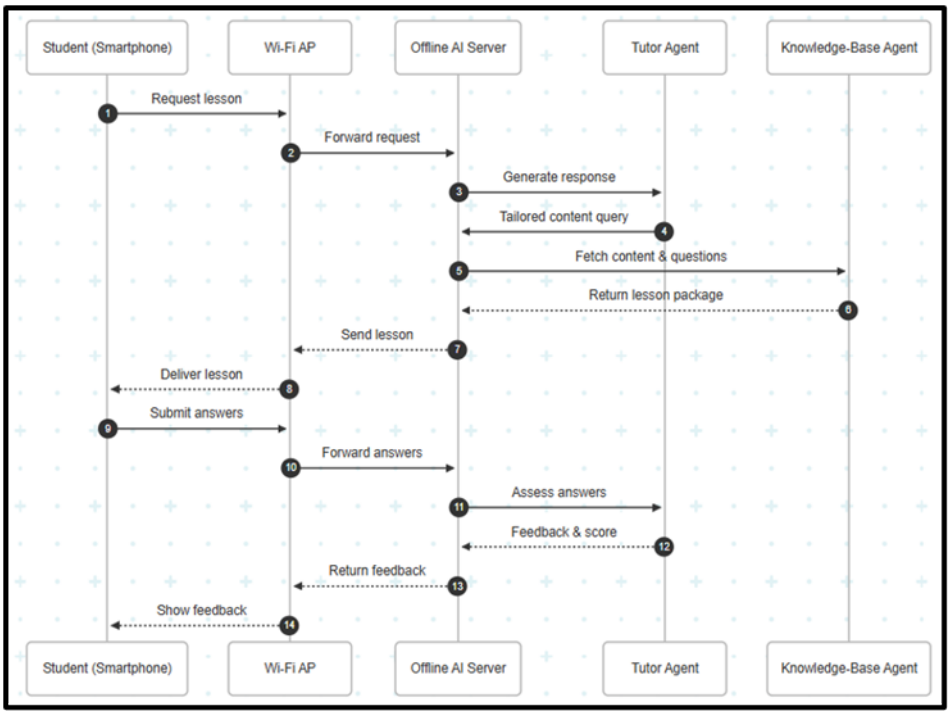
- Expand to lightweight AI apps on Android feature phones.
- Integrate solar-powered community learning hubs across Gaza and neighbouring regions.
- Localize NLP models for Arabic dialects; add speech interface for visually impaired students.
- Implement federated learning to improve models across multiple campuses securely.

- Seek funding and partnerships to scale pilots from 5,000 to 20,000 students within two years.
- Integrate Arabic GPT tutor modules after dataset localisation.
- Deploy agent-to-agent mesh synchronisation over local Wi-Fi.
- Formalise a shared knowledge-base schema (with ITU) for crowdsourced content.

## 3 Use Case Requirements

- **REQ-01**: It is of added value that we have access to low-power single-board computer (Raspberry Pi 4 or equivalent) with 512 GB SSD.
- **REQ-02**: It is of added value that we have access to solar panel + battery able to power server for ≥12 h blackout.
- **REQ-03**: It is of added value that we have access to local 2.4 GHz/5 GHz Wi-Fi AP with 30 m indoor range.
- **REQ-04**: It is of added value that we have access to quantised NLP and recommender models <200 MB.
- **REQ-05**: It is of added value that we have access to Arabic language support (tokeniser, embeddings, TTS/STT).
- **REQ-06**: It is of added value that we have access to AES-256 device encryption and local-only data storage.
- **REQ-07**: Each hub must run micro-agents exposed via REST/gRPC.
- **REQ-08**: Hubs must support LAN multicast or WebSocket for agent mesh sync.
- **REQ-09**: It is of added value that we have access to a knowledge-base stored as Markdown + vector embeddings (JSONL).

## 4 Sequence Diagram



## 5 References

[1] S. Y. Baroud and M. N. Qureshi, "Harnessing Industry 4.0: A Framework for Integrating AI and Emerging Technologies," in 2024 6th International Workshop on Artificial Intelligence and Education (WAIE), 2024.

[2] S. Y. Baroud and N. A. Yahaya, "RAPID : Integrating AI and Multi-Agent Systems for Enhanced Traffic Management Framework with," in International Conference on Innovation and Intelligence for Informatics, Computing, and Technologies (3ICT), 2024.

[3] S. Y. Baroud and N. A. Yahaya, "Optimizing Smart Manufacturing Through Multi-Agent Integration of Machine Learning Models : A conceptual framework," in 2024 RIVF International Conference on Computing and Communication Technologies, 2024.

[4] S. Y. Baroud, N. A. Yahaya, and A. M. Elzamly, "Cutting-Edge AI Approaches with MAS for PdM in Industry 4.0: Challenges and Future Directions," Journal of Applied Data Sciences, vol. 5, no. 2, pp. 455–473, 2024, doi: 10.47738/jads.v5i2.196.

# Use case -5: AI – Driven Virtual Instructor for 3D Virtual Electronics Lab

[Partner Organization - O-SKILLS PLC]

**Organization:** Addis Ababa Science and Technology University

**Country:** Ethiopia

**Contact Person(s)**:

Afrah Hussein Seid (afrahhussein04@gmail.com)

O-SKILLS PLC (info@oskills.org, sumeya.hussein@oskills.org)

## 1 Use Case Summary Table

| Category | Education |
|---|---|
| The problem to be addressed | Safety and cost limitations in physical electronics labs, gender/social bias in participation, unequal access to quality STEM education. |
| Key aspects of the solution | Education (Education Technology / Equity in STEM / Virtual Labs)<br>Real-time AI tutoring, adaptive task complexity, error diagnostics via circuit metadata, multilingual voice/text support, fairness-aware feedback. |
| Technology keywords | Virtual Labs, GNN, Transformer NLP, SPICE Simulation, Reinforcement Learning, Adaptive Learning, Multilingual ASR |
| Data availability | [2], [3] |
| Metadata (type of data) | Circuit diagrams (graph-based, .cir file), voice/text queries, student interaction logs (timestamped), adaptive feedback and performance scores |
| Model training and fine-tuning | GNNs trained on SPICE-netlist circuits; BERT fine-tuned for electronics Q&A in Amharic; RL agents trained on synthetic student personas |
| Testbeds or pilot deployments | Planned 9-month pilot at OSKILL PLC with anonymized data logging and confidence/skill improvement tracking |

## 2 Use Case Description

### 2.1 Description

The AI-Driven Virtual Instructor for 3D Virtual Electronics Lab is an innovative educational technology solution designed to improve equitable access to practical STEM education, particularly in resource-limited environments. This use case aims to replicate real-life electronics

lab experiences in a fully interactive 3D virtual environment, while integrating AI-powered instructional support to guide students through experiments, assess performance, and provide real-time, personalized feedback. The platform specifically addresses challenges such as the lack of physical lab infrastructure, shortage of expert instructors, language barriers in technical learning, and disparities in access to STEM resources. The AI instructor functions as a virtual mentor, capable of understanding and responding to voice commands in both Amharic and English, guiding learners through simulations, evaluating circuit designs, and offering adaptive hints based on student input via text, speech, or visual interactions.

Built using Unity, the 3D simulation engine models electronic components and behaviors, making the lab accessible from standard computing devices without the need for expensive VR headsets. The AI system includes a speech recognition module that converts spoken Amharic or English into text, and a natural language processing (NLP) and dialogue management system to interpret student intent and determine contextually appropriate AI responses. An AI evaluation module analyzes both circuit graphs and visual layouts to assess design accuracy, while Clarifai is used to manage multimodal data—such as images of circuits, voice inputs, and session metadata—for tagging and continuous learning system improvement. All student interactions are logged with metadata including error types and time-on-task, supporting adaptive feedback loops and reporting.

The expected outcomes of the project include significantly improved access to quality STEM education for students in rural or underfunded schools, increased student engagement and retention through interactive and hands-on learning experiences, and enhanced personalization of instruction based on real-time AI analysis. Moreover, the inclusion of Amharic speech recognition promotes linguistic inclusion, and the system enables data-driven insights into student learning outcomes, aiding teachers and institutions in making informed educational decisions.

Use Case Status: The use case is part of a large product development.

### Partners

O-SKILLS PLC, through its O-SKILLS Technical Education and Training Centre, serves as a key partner by leading the development of the 3D Virtual Electronics Lab environment. Their role involves designing and implementing the immersive 3D virtual lab space that accurately simulates electronics experiments. In parallel, the AI-driven instructional assistant—responsible for real-time guidance, speech recognition, and performance evaluation—will be developed by AASTU students. O-SKILLS will closely collaborate with the AI development team to ensure smooth integration of the AI instructor into the 3D lab, facilitating effective interaction between the virtual environment and the intelligent tutor. [1]

## 2.2 Benefits of use case

The project advances equitable access to technical education by replacing physical lab infrastructure with an AI-driven 3D virtual electronics lab, enabling underserved students to acquire job-relevant skills in electronics and digital literacy. Its adaptive learning paths and real-time feedback bridge educational disparities, ensuring inclusive skill development aligned with STEM employment demands. Multilingual support (e.g., Amharic) and gender-neutral design further dismantle barriers to participation, targeting systemic gaps in technical and vocational training.

By anonymizing interactions and employing culturally sensitive AI tutors with voice-based interfaces, the platform counters gender biases in STEM education. It fosters confidence and engagement among women and girls, addressing UNESCO-reported underrepresentation (<30% of global researchers) and societal barriers. Equal access to hands-on learning opportunities promotes empowerment and parity in STEM fields.

The system integrates AI and NLP to innovate education delivery, simulating advanced lab tools (e.g., oscilloscopes) at minimal cost through platforms like Unity/Unreal Engine. Its offline compatibility ensures accessibility in low-connectivity regions, reducing reliance on expensive infrastructure while accelerating adoption of scalable, AI-driven Education Tech solutions.

The virtual lab democratizes electronics education by removing geographic and socioeconomic barriers. Support for non-English languages and low-resource settings ensures marginalized groups—including rural, low-income, and female students—gain equitable access to practical training, leveling opportunities for participation in STEM-driven economies.

## 2.3 Future Work

The development of the AI-Driven Virtual Instructor for 3D Virtual Electronics Lab will progress through a structured, multi-phase roadmap to transition the concept from prototype to a scalable, impactful educational tool. Initial efforts will focus on constructing a proof-of-concept prototype, integrating foundational 3D circuit simulations with voice recognition to validate the AI instructor's feasibility in guiding students through basic electronics experiments. Concurrently, advanced AI capabilities will be developed, including multilingual natural language processing (NLP) for Amharic and English, adaptive feedback mechanisms, and machine learning models trained to evaluate circuit performance in real time. A robust 3D simulation engine, built using platforms such as Unity or Unreal Engine, will be engineered to replicate realistic electronics experiments, ensuring seamless interoperability with AI-driven guidance and assessment systems.

To maximize accessibility, user experience (UX) design will prioritize intuitive interfaces tailored to students in resource-constrained regions, while localization efforts will expand language support to additional Ethiopian working languages. Strategic collaboration with institutions like the O-SKILLS Technical Education and Training Center will ensure curriculum alignment with national and global STEM standards. Scalability will be addressed through cloud-based deployment, enabling high-concurrency access across diverse devices and minimizing hardware dependencies.

The AI's assessment framework will be refined to deliver personalized feedback and evaluate circuit accuracy, supported by continuous model training to enhance pedagogical effectiveness. Pilot testing in rural and underserved schools will quantify the system's impact on STEM proficiency, with iterative improvements informed by real-world data. Upon successful validation, the platform's scope will broaden to encompass physics and electrical courses, establishing a comprehensive virtual STEM ecosystem.

# 3 Use Case Requirements

- **REQ-01**: Real-Time Circuit Evaluation Using GNNs

  The system must utilize Graph Neural Networks (GNNs) to analyze and validate circuit connections in real-time, ensuring accuracy of student designs within the 3D lab.

- **REQ-02**: Transformer-Based NLP for Multilingual Support

  Transformer-based Natural Language Processing (NLP) models must be implemented to understand and respond to student queries in multiple languages, particularly Amharic and English.

- **REQ-03**: Speech Interface with ASR and TTS

  The AI instructor must include Automatic Speech Recognition (ASR) for converting spoken language into text, and Text-to-Speech (TTS) for delivering spoken feedback in local languages like Amharic.

- **REQ-04**: 3D Simulation Integration (Unity/Unreal Engine)

  The system must be developed in Unity or Unreal Engine to support an interactive, immersive, and scalable virtual electronics lab environment.

- **REQ-05**: Adaptive Learning with Reinforcement Learning

  The platform must incorporate a reinforcement learning engine to adapt instruction and difficulty levels based on student performance and learning patterns.

- **REQ-06**: Fairness Through Data Anonymization

  Tools must be integrated to anonymize sensitive demographic metadata in training and usage logs, ensuring fairness and compliance with ethical AI standards.

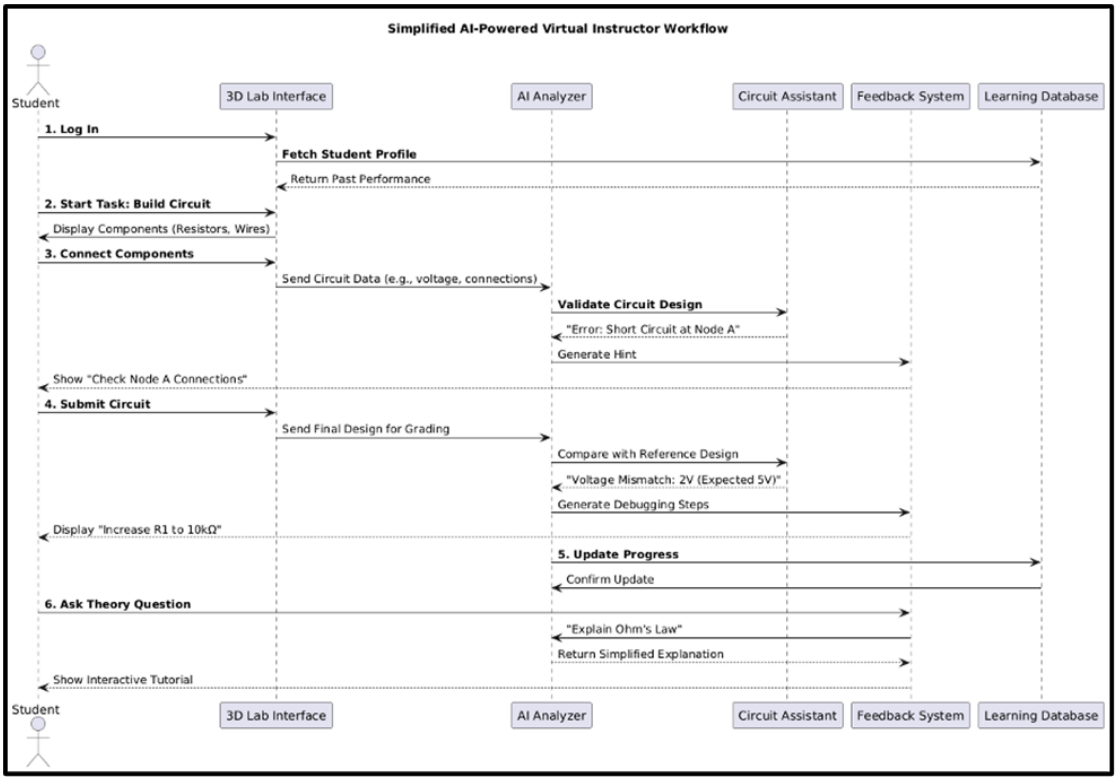
- **REQ-07**: Offline Inference Using Lightweight AI Models

  Quantized and optimized AI models must support offline inference on low-power or bandwidth-constrained devices to ensure usability in remote or under-resourced areas.

- **REQ-08**: Secure Performance Tracking and Visualization Dashboards

  The system must log educational performance data securely and provide dashboards for both students and educators to track progress, identify gaps, and support data-driven instruction.

## 4 Sequence Diagram




## 5 References

[1] https://oskills.org/visual-hub/

[2] https://github.com/DagsHub/3D-model-datasets - Dataset 1

[3] https://github.com/mingli-ai/Publicly-Available-Datasets-for-Educational-Data-Mining/ - Dataset 2

# Use case 6: AI-Powered Personalized Learning for Neurodivergent Students

**Country:** Ethiopia

**Organization:** Addis Ababa Science and Technology University

**Contact Person**:

Firaol Teshale Negera (firaolteshale@gmail.com)

## 1 Use Case Summary Table

| Category | Education |
|---|---|
| The problem to be addressed | Traditional classrooms fail to adapt to neurodivergent learners' cognitive and emotional needs, leading to disengagement, anxiety, and poor learning outcomes. |
| Key aspects of the solution | Education (Education Technology, Accessibility, Mental Health)<br>Real-time adaptive learning assistant using AI to detect engagement, stress levels, and customize content formats. Combines emotion detection, NLP-based content generation, and accessibility tools. |
| Technology keywords | NLP, Transfer Learning, RAG (Retrieval-Augmented Generation), Facial Expression Analysis, Sentiment Analysis, Adaptive Interfaces, Multimodal AI. |
| Data availability | Public datasets (Affect Net, SLDR, Kaggle Learning Data); private data collected with consent during pilots. |
| Metadata (type of data) | Eye tracking, facial expression data, learning interaction logs, emotion-labelled feedback. |
| Model training and fine-tuning | Fine-tuning pre-trained models for emotion recognition and adaptive content delivery; RAG for context-aware content personalization |
| Testbeds or pilot deployments | Planned pilot testing in special education classrooms with validation by educational psychologists and therapists |
| Code Repositories | Will be available upon completion of prototype on GitHub |

# 2 Use Case Description

## 2.1 Description

### AI-Powered Personalized Learning for Neurodivergent Students

Millions of neurodivergent students including those with autism, ADHD, dyslexia, and sensory processing disorders face learning barriers in traditional classroom settings. Standardized teaching methods often fail to accommodate their diverse cognitive styles, resulting in disengagement, stress, higher dropout rates, and reduced academic performance.

This use case proposes a real-time adaptive AI-powered learning assistant that supports neurodivergent learners by adjusting instructional content and delivery based on each student's emotional and cognitive state.

The system uses multimodal AI combining eye tracking, facial expression recognition, behavioral analysis, and sentiment detection to continuously assess engagement and cognitive load. Based on this input, the AI dynamically adjusts the pacing, format, and complexity of learning content using NLP-driven content generation and Retrieval-Augmented Generation (RAG). The system also provides accessibility support through features like text-to-speech, dyslexia-friendly fonts, and guided relaxation breaks when stress or fatigue is detected.

Unlike existing AI-based education tools that focus solely on academic performance, this system prioritizes emotional awareness and mental well-being. It not only customizes the learning experience but also helps students manage cognitive stress, thereby improving retention, motivation, and overall educational equity.

Expert validation is a core part of the approach. Educators and psychologists will evaluate AI-generated outputs during pilot phases, ensuring reliable detection of frustration, disengagement, and overload. Furthermore, the system supports personalization over time by learning each student's behavioral and emotional baseline.

Privacy and ethical design are embedded into the system. Emotion and biometric data are collected only with informed consent, processed locally where possible, and secured using anonymization techniques and privacy-by-design standards.

### Partners

At present, this project does not have formal institutional or organizational partnerships. However, the future implementation roadmap includes outreach to:

- Special education schools and therapy centers for real-world pilot testing.
- Universities and AI research groups for model development and validation.
- Mental health professionals and psychologists to help co-design stress detection mechanisms and ensure ethical oversight.

The project is designed with openness in mind, aiming for future collaboration with stakeholders in both education and mental health sectors, as well as open-source AI communities.

## 2.2 Benefits of use case

This use case directly delivers inclusive, adaptive learning experiences tailored to the unique needs of neurodivergent students. By leveraging AI to adjust content in real-time based on emotional and cognitive feedback, the system ensures that each learner receives an education aligned with their learning style and mental well-being.

It addresses a population often left behind by standardized education systems. Neurodivergent students frequently face academic and social marginalization due to a lack of individualized support. This AI solution helps level the playing field by offering dynamic assistance and equal access to digital education tools.

Furthermore, it monitors cognitive overload, emotional fatigue, and stress. The system recommends supportive interventions such as mindfulness breaks or slower pacing to prevent burnout, anxiety, or learning fatigue. By focusing not only on academic progress but also on mental health, the project promotes a holistic approach to student development.

## 2.3 Future Work

1. Data Collection & Expansion

The project will involve collecting diverse, ethically sourced multimodal data from neurodivergent learners. This includes facial expression cues, eye movement patterns, engagement signals, and behavioral interactions across various educational settings. All data collection efforts will prioritize informed consent, privacy, and student well-being.

2. Proof of Concept Development

A minimum viable prototype will be developed to demonstrate real-time emotional assessment and dynamic content adaptation. Initial deployments will take place in collaboration with special education institutions to validate responsiveness and accuracy in real-world classroom environments.

3. Model Development

Pre-trained AI models for sentiment analysis, attention tracking, and adaptive content generation will be fine-tuned using neurodivergent-focused datasets. Retrieval-Augmented Generation (RAG) will be incorporated to align content with curriculum standards and personalized learning contexts.

4. Variations and Extensions

Planned future extensions of the system include:

- Dashboards for educators to track collective emotional and cognitive trends
- Integration with assistive or wearable technologies
- Embedded mindfulness tools (e.g., breathing cues, calming visuals, adaptive audio) tailored for sensory-sensitive learners

5. Standards Development

Efforts will be made to contribute to emerging global standards on AI in neurodivergent education. This includes aligning with ITU's architectural frameworks (such as Y.3172), and promoting fairness, transparency, and responsible emotional data usage.

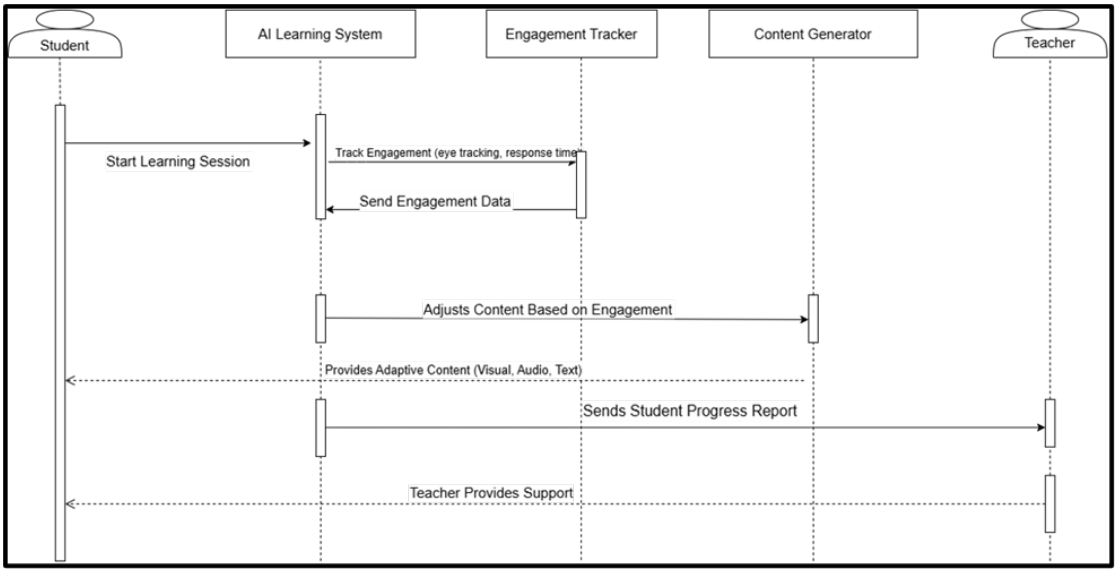
6. Reference Tools and Simulation Environments

Interactive simulation environments will be developed to allow researchers, educators, and developers to test adaptive learning flows, emotional response models, and feedback loops under controlled experimental conditions. These tools will be open source to encourage community collaboration and iterative improvement.

## 3 Use Case Requirements

| Requirement ID | Description |
|---|---|
| REQ-01 | The system shall detect and classify emotional states (e.g., stress, fatigue, frustration) using real-time facial expression and behavioral data. |
| REQ-02 | The system shall generate adaptive learning content in multiple formats (text, visual, audio) based on user preference and engagement level. |
| REQ-03 | The AI tutor shall recommend micro-interventions (e.g., breaks, breathing exercises) when cognitive overload is detected. |
| REQ-04 | The system shall allow personalized calibration for students with unique emotional expression baselines. |
| REQ-05 | All collected data shall be anonymized and processed with consent, following strict privacy-by-design principles. |
| REQ-06 | The system shall support on-device processing of sensitive data (e.g., facial images, eye-tracking) wherever possible to enhance privacy. |
| REQ-07 | The system shall provide interpretable outputs and real-time feedback to educators and parents. |
| REQ-08 | The model shall support fine-tuning with neurodivergent-specific datasets and use RAG for curriculum-aware content personalization. |
| REQ-09 | The system shall log activity and emotional feedback for longitudinal learning progress tracking. |

## 4 Sequence Diagram




## 5 References

[1] AI for Students with Learning Disabilities: A Systematic Review – ResearchGate https://www.researchgate.net/publication/374295269_AI_for_Students_with_Learning_Disabilities_A_Systematic_Review

[2] Emotion Recognition and Cognitive Load Detection in Learning Environments – CORE.ac.uk https://core.ac.uk/reader/29584880

[3] Ali Mollahosseini, Behzad Hasani, Mohammad H. Mahoor. AffectNet: A Database for Facial Expression, Valence, and Arousal Computing in the Wild. IEEE Transactions on Affective Computing. 2019. https://doi.org/10.1109/TAFFC.2017.2740923

[4] Pavel Grill, Jana Tučková. Speech Databases of Typical Children and Children with SLI. PLoS ONE. 2016. https://doi.org/10.1371/journal.pone.0150365

[5] Angel Olider Rojas Vistorte, Angel Deroncele-Acosta, Juan Luis Martín Ayala, Angel Barrasa, Caridad López-Granero, Mariacarla Martí-González. Integrating Artificial Intelligence to Assess Emotions in Learning Environments: A Systematic Literature Review. Frontiers in Psychology. 2024.

## Use case 7: UN Women AI School

**Country:** Thailand

**Organization:** UN Women Regional Office for Asia and the Pacific

**Contact Person(s)**:

Emad Karim (emad.karim@unwomen.org)

Federica Tornincasa (federica.tornincasa@unwomen.org)

### 1 Use Case Summary Table

| Category | Education |
|---|---|
| The problem to be addressed | AI has the power to accelerate progress toward gender equality, but only if women and feminist leaders are equipped with the knowledge, skills, resources and enabling environments to actively shape its development and use. |
| Key aspects of the solution | Artificial Intelligence<br>• Custom GPTs developed with and for participants to enhance productivity and knowledge management<br>• Innovation hub to promote practical AI solution to accelerate progress toward gender equality. |
| Technology keywords | Generative AI, Custom GPTs, Prompt Engineering, LLMs |
| Data availability | Public (Training materials and custom GPTs generated for the AI School will be available on UN Women's website). |
| Metadata (type of data) | Text data |
| Model training and fine-tuning | The UN Women AI School introduces participants to cutting-edge Generative AI (GenAI) technologies and practical approaches tailored for diverse user needs. Participants explore tools such as ChatGPT, Gemini, Notebook LLM, and Microsoft Copilot, applying them to use cases in research, knowledge production, content creation, advocacy, and campaign design. The School emphasizes prompt engineering, responsible AI use, and no-code custom GPT development. Participants learn how to adapt AI solutions to their workflows—creating policy briefs, donor proposals, knowledge products, and advocacy materials—without requiring advanced technical skills. The curriculum integrates hands-on learning, real-world case studies (e.g., UN Women's Humanitarian GPT), and a focus on inclusive, gender-responsive AI applications. The program also promotes agile learning design, enabling continuous updates aligned with the fast-paced evolution of AI technologies. |
| Testbeds or pilot deployments | [3], [4], [5] |

# 2 Use Case Description

## 2.1 Description

AI has the power to accelerate progress toward gender equality—but only if women and feminist leaders actively shape its development and use. From tackling online gender-based violence to ensuring that AI-powered tools are free from bias, feminist voices must be central in AI innovation. The UN Women AI School is designed to bridge the AI knowledge gap and empower participants to use AI for advocacy, campaign strategy, and decision-making.

Building on the lessons learnt in its first edition (2024), the 2025 UN Women AI School edition will enable UN staff, policymakers and civil society to leverage AI for more effective advocacy, program implementation, and organizational transformation. The Academy will have three different tiers, customized to participants' needs:

- Public Cohort: including women, young people and feminist leaders, the Public Cohort will have a 3-month duration and an extensive programme aimed at providing foundational knowledge on ethical and inclusive AI and to build participants' capacity to develop AI-powered solutions to urgent global challenges, including Climate Justice, Tackling Technology-Facilitated Gender-Based Violence (TFGBV), and Humanitarian Action.
- Policy-Makers Intensive: targeted at policy and decision makers, this tier will equip them with foundational knowledge and skills for regulatory design and public-service deployment
- UN Staff Bootcamp: targeted at staff members of UN Women an other UN agencies, this tier will provide participants with UN-specific examples, leveraging in-person labs to build immediate, workstream-ready prototypes

In its pilot project in 2024, the UN Women AI School successfully trained more than 1,000 UN Women Staff, UNCTs, young leaders, and CSOs, achieving a 100% satisfaction rate. This demonstrated the program's effectiveness in delivering high-quality, gender-responsive AI education at scale. The 2025 target is to reach 1,200 UN Staff and young leaders working on AI for Gender Equality, further expanding its impact and reach.

By partnering with governments, private sector entities, and educational institutions, this initiative scales AI training across communities, ensuring that women, youth, and persons with disabilities are empowered to thrive in the AI-driven economy. The program fosters inclusive AI development, ensuring responsible AI governance aligned with ITU and UN AI ethics principles.

**Partners**

- UN Women (Secretariat and Manager)

UN Women serves as the lead agency, managing the overall coordination, design, and implementation of the AI School. The organization ensures alignment with the UN Women mandate, UN 2.0, and broader global digital and gender equality strategies.

- 30 for 2030 Youth Leadership Network

As co-designers and facilitators, members of the 30 for 2030 Youth Leadership Network actively contribute to curriculum development, learning facilitation, and peer mentoring. They also

serve as members of the AI School Steering Committee, ensuring youth leadership and diverse perspectives in program governance.

- UN Issue-Based Coalition (IBC) on Digital Transformation

The UN IBC promotes cross-agency collaboration, ensuring that the AI School aligns with regional digital transformation strategies and supports collective UN system efforts. Support AI School UN Staff Track.

- AI School Steering Committee (Independent and Voluntary Group)

The Steering Committee guides and oversees the AI School's design, implementation, and evaluation to ensure the program remains inclusive, representative, and impactful. Members include UN Women country office staff, experts from UN agencies (including UNICEF, UNDP, and ESCAP), and independent specialists in gender, technology, and education.

## 2.2 Benefits of use case

Women remain underrepresented in STEM and AI sectors. This initiative empowers women with AI-driven education, increasing workforce participation and leadership in technology fields.

By equipping women and marginalized groups with AI skills, this initiative enhances workforce readiness, improves employment outcomes, and drives inclusive economic growth.

AI-powered education includes training on AI applications for climate resilience, supporting women-led innovation in sustainable technologies and green jobs.

## 2.3 Future Work

To further scale and sustain the impact of this initiative, we propose the following future work:

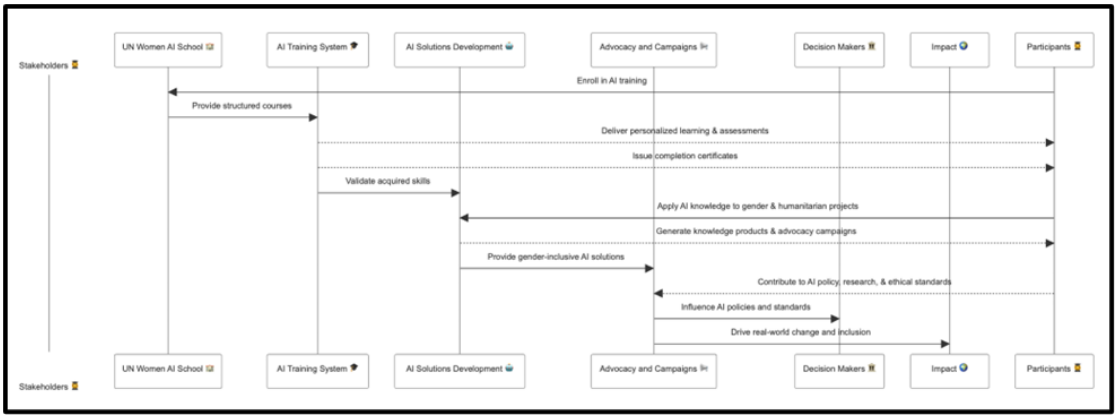
- Data Collection & Insights Development: Expand sex-disaggregated AI data research to identify barriers and opportunities for women in AI careers in the Asia-Pacific region, both at the national and the regional level. Strengthen AI-driven labor market analytics to tailor skill development programs based on emerging workforce trends.
- Proof of Concept Development: Enhance UN and civil society organizations workflows with AI productivity tools. These will include custom GPTs trained with relevant sources (including UN Reports and datasets from National Data Institutions) and adopting the Human Rights Based Approach and a Gender Equality, Disability and Social Inclusion (GEDSI) lens.
- Design of Gender Equality Datasets & Exploration of Gender-Inclusive LLMs: Develop comprehensive datasets that ensure gender-equitable AI training, reducing bias in AI systems by applying GEDSI lenses and disaggregating data by gender, SOGIESC, disability, and socioeconomic status. Explore the development of gender-inclusive Large Language Models (LLMs) to create fairer and more representative AI applications.
- New Variations & Extensions: Extend AI skilling to include AI for social impact, AI ethics, and AI entrepreneurship programs, fostering women-led AI startups.
- Standards Development: Collaborate with ITU to develop AI governance standards that prioritize gender inclusivity, ethical AI, and workforce equity.
- Reference Tools & Simulations: Develop interactive AI-powered simulations for AI ethics training, ensuring inclusive AI development practices are embedded in industry norms.

By scaling this initiative, we ensure that AI's economic opportunities are equitably distributed, particularly benefiting women and marginalized groups in emerging markets.

## 3 Use Case Requirements

- **REQ-01**: It is required for users to register and complete the authentication process.
- **REQ-02**: The system should support various student inputs, including handwritten image uploads, typed text, and scanned paper submissions.
- **REQ-03**: The system should support real-time processing capabilities in order to deliver immediate feedback to students and teachers.
- **REQ-04**: The system should support educator-side interfaces for reviewing AI grading results and providing manual overrides when necessary.

## 4 Sequence Diagram




## 5 References

[1] https://asiapacific.unwomen.org/en/stories/announcement/2025/03/un-women-ai-school-opens-for-changemakers

[2] https://asiapacific.unwomen.org/en/stories/feature-story/2025/01/young-leaders-and-un-women-asia-and-the-pacific

[3] https://chatgpt.com/g/g-EzjBD9zmS-unw-humanitarian-hub

[4] https://chatgpt.com/g/g-LRLaoTwrk-tfgbv

[5] https://chatgpt.com/g/g-FTqlfJl7d-weps-advisor

## Use case 8: Little Signs

**Country:** Zimbabwe

**Organization:** Sign<Tech>

**Contact Person(s)**:

Jamain Vimbanashe Chindoore (tatendalerrical@gmail.com)

Cresentia B Moyo (moyobcresentia@gmail.com)

## 1 Use Case Summary Table

| Category | Education |
|---|---|
| The problem to be addressed | The development of a real-time speech-to-sign-language translator addresses critical communication barriers faced by the deaf and hard-of-hearing communities. These barriers hinder access to vital information and limit participation in social, educational, and professional environments. A further challenge lies in ensuring that such a system can accurately translate not only the linguistic content but also the emotional undertones, including manipulative tactics like emotional blackmail, that might arise in conversations. Emotional blackmail, which leverages fear, obligation, and guilt to coerce behavior, must be identified and appropriately conveyed in the translation process to empower users and ensure emotional integrity in communication. This highlights the need for advanced natural language processing and gesture-recognition technologies capable of capturing both literal meaning and emotional nuance. |
| Key aspects of the solution | Inclusive Education Technology<br>• Speech recognition modules (speech is converted to text via the use of a an AI deep learning model through Google Speech API)<br>• Natural Language Processing (this aspect include the processing and understanding of grammar, intent, and sentence structure which is in alignment with the Zimbabwe sign language. In other words, deaf culture using GPT for text parsing and translation into sign compatible structures.<br>• Sign Language Mapping Engine (this matches translated text to correct sign gestures )<br>• Gesture and Avatar Animation Module (this displays the signs visually via a 3D avatar )<br>• Zimbabwean Sign Language Gesture Database<br>• Real-time processing system<br>• Accessibility & Localization Layer (using AI to customize translation to local sign language and cultural context) |

(continued)

| Category | Education |
|---|---|
| Technology keywords | Speech Recognition, Natural Language Processing (NLP), Text-to-Sign Translation, Sign Language Animation, Real-Time Processing, 3D Avatar Rendering, Machine Learning (ML), Deep Learning, Sequence-to-Sequence Model, Transformer Models, Speech-to-Text Engine, Gesture Mapping, Sign Language Grammar, Multimodal AI, Contextual Translation, Avatar Animation Pipeline, Sign Language Dataset, Pose Estimation, Low-Latency Inference, Human-Computer Interaction (HCI), Semantic Parsing, Language Modeling, Phoneme Recognition, Audio Signal Processing, Accessibility Technology, Inclusive Education Tools |
| Data availability | • ASL Database for training the system models: [5]<br>• Dataset 2: [6] |
| Metadata (type of data) | Text, real-time audio, visual, images |
| Model training and fine-tuning | In this real-time speech-to-sign language translation system, artificial intelligence and machine learning technologies are integrated to facilitate seamless communication. The process begins with automatic speech recognition (ASR) models, such as Google Speech, which transcribe spoken language into text. This textual data is then processed using natural language processing (NLP) techniques, employing the BERT & T5 models, to understand context and structure. Subsequently, sequence-to-sequence models, often based on transformer architectures, translate the processed text into sign language glosses. These guide the generation of corresponding sign gestures, which are animated using 3D avatars. The avatars' movements are driven by pose estimation and animation models trained on sign language datasets, ensuring accurate and culturally appropriate representations. To optimize performance for real-time applications, the system undergoes fine-tuning through transfer learning and employs techniques like model pruning and quantization, enhancing efficiency and responsiveness all with the use of an internet connection. |
| Testbeds or pilot deployments | www.littlesigns.co.zw<br>King Gorge The 4th Secondary school |

# 2 Use Case Description

## 2.1 Description

Communication barriers between deaf individuals and hearing communities remain a significant challenge worldwide. Spoken languages often exclude individuals who rely on sign language for communication, limiting access to essential services, education, employment opportunities, and social interactions. This challenge is amplified in environments requiring rapid exchange of information, such as classrooms, workplaces, or emergencies. Real-time speech-to-sign-language translators aim to bridge this gap, promoting inclusivity and fostering a sense of belonging. Moreover, the philosophy that "disability is not an inability" reinforces the importance of creating tools that enable individuals to fully express their potential without limitations imposed by communication barriers.

The primary aim of the real-time speech-to-sign-language translator is to develop a seamless communication system for the deaf community, enhancing accessibility and participation in various spheres of life. Objectives include:

- Ensuring accurate translation of spoken language to sign language in real-time.
- Capturing the emotional and contextual nuances of speech for more meaningful communication.
- Incorporating flexibility for diverse sign languages, accents, and cultural expressions.
- Empowering individuals to engage equally in educational, professional, and social spaces.

### Innovative Technological Approach

1. **Advanced Speech Recognition:** AI models trained to recognize diverse accents, languages, and speech patterns in real-time.

2. **Gesture Recognition and Generation:** Neural networks process gestures and facial expressions to create visually accurate sign language translations.

3. **Machine Learning Adaptability:** Models that learn and improve with user feedback, enabling customization for personal or cultural variations in sign language.

4. **Integration with Wearable Devices:** Smart glasses or portable displays can discreetly provide sign language output, ensuring accessibility in any environment.

5. **Ethical and Inclusive Design:** Collaboration with deaf communities ensures cultural relevance and respect for sign language nuances.

Expected Impact and Benefits:

- Promote Social Inclusion: Enable seamless interaction between hearing and deaf individuals, eliminating barriers that hinder collaboration and understanding.
- Increase Accessibility: Provide equal opportunities in education, employment, and public services, empowering individuals with disabilities to thrive.
- Enhance Independence: Allow deaf individuals to navigate diverse environments independently, without requiring constant interpreter support.
- Foster Innovation: Push technological boundaries, leading to advancements in AI, machine learning, and assistive technologies.
- Celebrate Diversity: Embrace the idea that "disability is not an inability," showcasing the capabilities and contributions of individuals with disabilities.

### Partners

- National Disability Association of Zimbabwe: they are helping us with giving us access to the deaf community so we can learn the Zimbabwe sign language.

## 2.2 Benefits of use case

Our project (Little Signs) enables the deaf and hard-of-hearing students to access spoken classroom content in real time while fostering and promoting an inclusive learning environment for all from preschool level up to tertiary level. It also supports equal participation.

Our project reduces communication-related stress and social isolation among the deaf learners. It promotes mental well-being by enabling active participation and understanding in social and learning contexts.

Little Signs uses AI, speech recognition, and 3D animations to solve a real world problem while contributing to the development of smart educational and communication infrastructure.

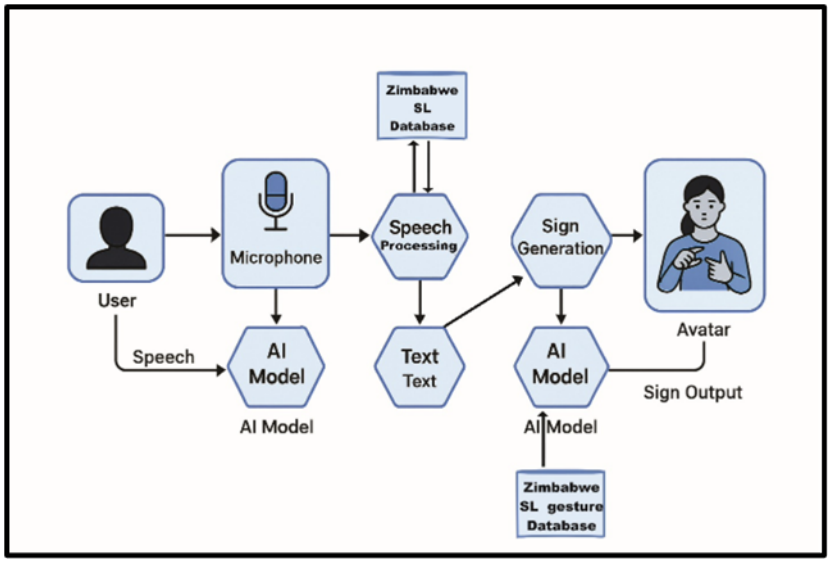
Little Signs helps bridge the communication gap between hearing and non-hearing individuals in the educational system.

- King George the 4th Secondary School : They are our primary pilot school as they offer education to the population of kids with disabilities .
- Potraz: is providing an infrastructure to build and test the system and make sure it is a success
- 2.3 Future Work
- We are planning to then increase the sign language grammar and database size in the future so as to cover a wider African population
- We are also planning on improving our NLP models to be able to choose the correct grammar and and technical terms for different subjects.
- High-performance processors and GPUs for real-time processing.
- Cameras with high frame rates and resolution to accurately track hand and facial movements for sign generation.
- Advanced real-time audio recognition and Natural Language Processing (NLP) systems.
- Extensive video datasets of various sign languages
- Data from real-world scenarios with users of various accents and speech patterns.

## 3 Use Case Requirements

- **REQ 1**: It is critical that we have Zimbabwean Sign Language Database, grammar and gesture database
- **REQ 2**: It is critical that we have Hardware including external microphones, high-end GPUs, high end CPUs, projectors, and projector screens
- **REQ 3**: It is critical that we have a linguist specializing in Sign Language
- **REQ 4**: It is critical that we have human resources
- **REQ 5**: It is critical that we have internet connectivity

## 4 Sequence Diagram




## 5 References

[1] Real-time Conversion of Sign Language to Text and Speech, and Vice Versa:

[2] https://jetir.org/papers/JETIR2311332.pdf

[3] Audio to Sign Language Using NLTK:

[4] https://www.jetir.org/papers/JETIR2306484.pdf

[5] A Survey Paper on Speech/Audio to Sign Language Translator for Deaf People:

[6] https://www.jetir.org/papers/JETIR2112300.pdf

[7] AfriSign: Machine Translation for African Sign Languages

[8] You can locate this paper on academic platforms like Google Scholar by searching the title "AfriSign: Machine Translation for African Sign Languages" or through repositories like ZORA (Zurich Open Repository and Archive).

[9] American Sign Language Linguistic Research Project (ASLLRP) ASL Sign Bank: http://dai.cs.rutgers.edu/dai/s/signbank

[10] How2Sign Dataset: https://how2sign.github.io/

## Use case -9: Using AI in Student Creative Art Translations and STEM Workshop

**Country:** Nigeria

**Organization:** Rising Hope Girls Educational Foundation (RHGE) (NGO)

**Contact Person(s)**:

David Ahanor (ahanordavid@gmail.com) - Digital Program Coordinator

Oghogho Ahanor (support@risinghopegirlseducation.com)

OlaKunle Festus (kunlefes089@gmail.com)

## 1 Use Case Summary Table

| Category | Education |
|---|---|
| The problem to be addressed | Limited access to digital tools and creative tools in learning in underserved communities in Nigeria has led to 60 percent of girls population with no access to digital STEM learning access. Poor undeveloped technological use skills of teachers, where 60 percent of their population are digitally illiterate and the low infrastructural system deficiencies has stumped innovative and creative use STEM techniques and tools in underserved digital areas. |
| Key aspects of the solution | Digital Learning, Education, AI in Creativity, STEM Empowerment, Digital Inclusion, Literacy, AI in STEM, Evidence-Based Learning<br>AI-generated translations of student art into digital formats, STEM learning, integration of AI tools (e.g., image recognition, generative models), visual storytelling, AI in NFTs, creative art work, evidence-based learning, coding, EdTech application |
| Technology keywords | Generative AI, Computer Vision, Digital Literacy, Machine Learning, AI Art, NLP, STEM Education |
| Data availability | Private dataset (student art images, STEM activity logs) |
| Metadata (type of data) | Visual (images), text (annotations), audio (optional narratives) |
| Model training and fine-tuning | Uses pre-trained generative models (ChatGPT, Copilot, DALL-E, Stable Diffusion, Openart.ai), prompt-tuning for contextual relevant |
| Testbeds or pilot deployments | Pilot projects conducted in schools across Kubwa, Dei Dei, Dutse, Byazhin regions in Abuja [12] [13] |

# 2 Use Case Description

## 2.1 Description

The poor access to digital tools and creative tools in learning in underserved communities in Nigeria has led to 60 percent of girls population with no access to Digital STEM learning access as reported by UNICEF. Poor undeveloped technological use skills of teachers, with 60 percent digitally illiterate and infrastructural system deficiencies has stumped innovative and creativity in STEM learning. All these issues have made evidenced literacy-based training in many public schools fall short in supporting educational STEM learning, digital illiteracy and exclusion of girls' especially in public schools and digitally unserved communities.

This use case demonstrates how the Rising Hope Girls Educational Foundation (RHGE) is utilizing and integrating GENAI to enhance creative expression and STEM learning for girls in underserved digital communities in Nigeria and support innovative use of AI in fostering girls' digital inclusion in digital underserved communities, promote gender inclusion campaigns, engage new cognitive literacy learning action in the use of EdTech workflows in STEM and STEAM with AI use cases in programs. The initiative integrates student-generated art with AI models in Project Based Learning solutions that is capable of translating their artworks into dynamic digital forms, such as visual animations, NFTs, AI-generated narratives, and immersive presentations. This has also broadened our other STEM workshops practices adoption in employing AI tools to help students simulate experiments, visualize scientific concepts, support AI creative workshop in evidenced based practice with other use applications, coding, digital learning solutions and explore careers in emerging tech fields.

In this work the Generative AI not only enhances the aesthetic value of creative outputs but also acts as a bridge to computational thinking, problem-solving, digital literacy development and innovation for students. It has shown to support CS use learning opportunities to stimulate girls' interest in STEM careers and drive arts creativity. It has also shown to support social campaign for girls in Challenge based learning adoption and in Project Based Evidence learning in the communities.

Language of instruction was English, as students engaged with English for instructional learning activities.

**Use case status:**

Active – ongoing deployments in community-based schools in 4 regions of North central Nigeria in Abuja FCT area of Dutse, Kubwa, Bhazhin, Dei Dei, where 6 schools have been approved for digital training by the NGO in supporting AI learning in STEM. The learning impact would support more than 1000 teachers and 7000 students in AI cognitive training. Also, the Use case was deployed in Challenge based learning solutions for students in the Ciena Solution Challenge by Digital Promise a US educational NGO. Our project use case was part of the 42 projects that was awarded with over 135 projects submissions. [9] [10]

**Partners**

- Schools in target regions: Hosting workshops and pilot programs
- Team4Tech: US based NGO, support for EBP resource in AI learning and GENAI implementation learning, Digital Literacy and Fluency development.

- Digital Promise: US based NGO. Support for CIENA solution in Student AI use in PBL
- TechSoup: Technical support partners for digital applications and access training tools
- SEB: Secondary education Board of Nigeria
- UBEB: Universal Educational Board of Nigeria
- NITDA: National Information Technological Agencies of Nigeria-NDLF
- Community Youth Volunteers: Local facilitation and student engagement

## 2.2 Benefits of use case

This has supported the learning and digital inclusion of girls in supporting cognitive use of AI tools to drive literacy and fluency skills in different digital applications for educational learning and digital inclusion, which provides quality education for learners especially in underserved digital environments and also maintain local literacy development targets of the Nigerian Digital Literacy Framework (NDLF).

The use case has supported girls' campaign in rights use initiatives in challenge-based learning (CBL) for rights advocacy for girls in the regions as adopted in the CIENNA Challenge of Digital Promise, where girls' rights were canvassed with

GENAI NFTs of Students creative artworks, which supports gender rights and promotion of girl's rights. (See student created blog link with their AI outputs for gender rights awareness and digital inclusion) [11]

The use of GenAI tools innovatively in EdTech integration to learning has seen digitally unserved regions exposed to evidenced based practice that supports educational learning infrastructures with the addition of these GENAI in the learning ecosystem and support internet/computer device support for STEM learning in local unserved digital schools

Finally, the workflows in art creation and digital STEM applications have been used in remote places and communities has seen that AI can drive ubiquitous learning opportunities to disadvantages learners and communities in supporting SDG 10.0 in reducing inequalities

- Community Youth Volunteers: Local facilitation and student engagement

## 2.3 Future Work

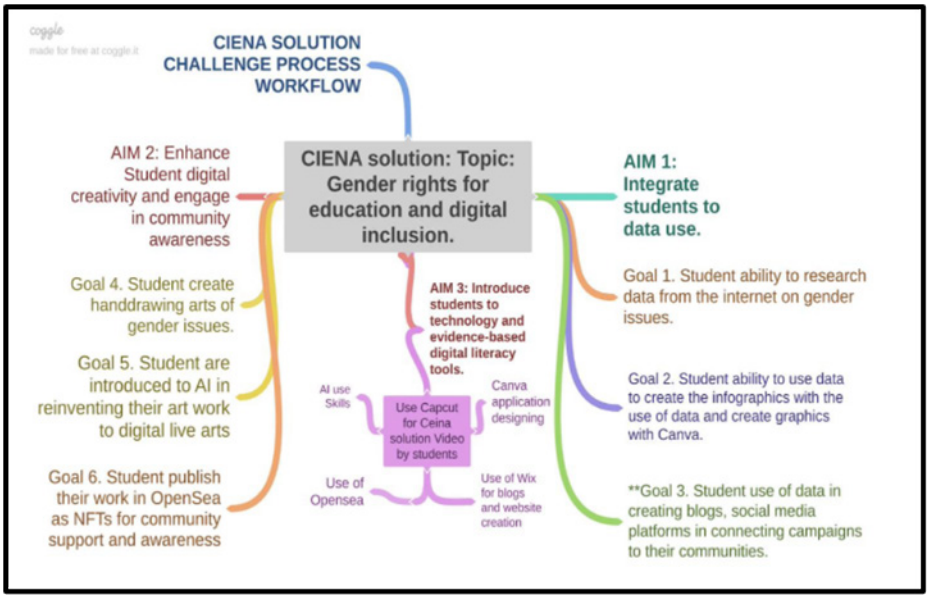
- Expand to more schools across northern Nigeria (see approval letters further training in schools)
- Develop local language voice-to-text AI translations for art storytelling to support for rural student that have challenges in language literacy and use of English instructions in AI use case
- Open-source creative outputs and training data
- Strengthen partnerships with global AI initiatives

# 3 Use Case Requirements

- **REQ-01**: It is critical that we have access to internet-enabled devices for running AI tools
- **REQ-02**: It is critical that we have consent and ethical data handling policies for student-generated content

- **REQ-03**: It is critical that we have integration of AI interfaces with low-bandwidth usability
- **REQ-04**: It is critical that we have visual and audio output rendering systems
- **REQ-05**: It is critical that we have teacher training on using AI-powered creative tools

## 4 Sequence Diagram




## 5 References

[1] ALIGN Programme. (2024). Whose hands on our education? Advancing learning and innovation on gender norms. ALIGN. Retrieved from https://www.alignplatform.org Britten, J. (2024). AI responsible use guidelines. Team4Tech. [PDF file].

[2] Britten, J. (2024). Rubric for selecting AI tools. Team4Tech. [PDF file].

[3] Brynjolfsson, E., & McAfee, A. (2014). The second machine age: Work, progress, and prosperity in a time of brilliant technologies. W.W. Norton & Company.

[4] Holmes, W., Bialik, M., & Fadel, C. (2019). Artificial intelligence in education: Promises and implications for teaching and learning. Center for Curriculum Redesign.

[5] Oyelere, S. S., Sanusi, I. T., Agbo, F. J., Oyelere, A. S., Omidiora, J. O., Adewumi, A. E., & Ogbebor, C. (2022). Artificial intelligence in African schools: Towards a contextualized approach. IEEE Xplore. Retrieved from https://ieeexplore.ieee.org/abstract/document/9766550

[6] OpenAI. (2024). Teaching with AI: Practical strategies for educators. Retrieved from https://openai.com

OECD. (2021). Artificial intelligence in society. OECD Publishing. https://doi.org/10.1787/5dc77ac0-en

[7] Romero, M., & Urmeneta, A. (2024). Creative applications of artificial intelligence in education.

[8] UNESCO. (2024). Generative AI in low- and middle-income countries: Potential and challenges. UNESCO. Retrieved from https://www.unesco.org/digital-learning-week

[9] https://digitalpromise.org/2025/04/16/announcing-the-2025-ciena-solutions-challenge-sustain ability-awards/ - Evaluation by Ciena Challenge about this project

[10] https://www.cienachallenge.org/en/custom/projects/view/26710 - Evaluation by Ciena Challenge about this project

[11] https://oghoghoahanorrhgf.wixsite.com/rhgecienastudentblog - Blog created by students

[12] https://risinghopegirlseducation.com/gender-rights-for-education-and-digital-inclusion-ciena-solution-activity-project/ - Pilot deployment

[13] https://drive.google.com/drive/folders/1Z7p1umOpUpB19KRoJxUgSQco0qw30ou6 - Examples of outputs from the programs and workshops

# Use Case 10: AI-Driven Personalization for Sustainable Development: Transforming Education, Wealth, Nutrition, and Health

**Country:** Uganda

**Organization:** Rein4ce Learning (U) Ltd.

**Contact Person(s)**:

Primary contact: James Boogere (jboogere@gmail.com)

Secondary contact: Kassim Wasswa (wasswakasim.wk@gmail.com)

## 1 Use Case Summary Table

| Category | Education |
|---|---|
| The problem to be addressed | To integrate AI into personalised, sustainable development |
| Key aspects of the solution | Life-long Learning<br>I. Optimize agricultural practice for institutional meal programs through AI-based climate analytics and resource management.<br>II. Monitor and manage learners' nutritional needs via smart wearables and AI-driven dietary analysis.<br>III. Provide personalized education through adaptive learning platforms, integrating health longevity training and nutrition data. |
| Technology keywords | AI (machine learning, Natural Language Processing (NLP), Internet of Things (IoT), cloud computing), data integration, wearable devices, and adaptive learning platforms |
| Data availability | Private |
| Metadata (type of data) | The metadata will vary based on the data type—educational progress, health indicators, nutrition data, and agricultural inputs. It includes:<br>• Timestamps (e.g. time of activity or data collection)<br>• Device IDs (e.g. unique identifiers for wearables/sensors), of Data)<br>• User demographics (e.g. age group, institution, region)<br>• Contextual information (e.g. environmental readings or academic session)<br>• Data source type (sensor, manual input, app logs, etc.) |

(continued)

| Category | Education |
|---|---|
| Model training and fine-tuning | Several types of AI models will be implemented depending on the data domain:<br>A) Education (Adaptive Learning)<br>• Model Type: Personalized learning recommendation engines using reinforcement learning, collaborative filtering, and Bayesian knowledge tracing.<br>• Goal: Recommend lessons, modules, or learning paths tailored to each learner's pace, needs, and performance.<br>B) Health & Nutrition<br>• Model Type: Time-series models, classification models, and decision trees (e.g., Random Forests, LSTM networks for temporal health patterns).<br>• Goal: Alert users/institutions on nutrition risks or recommend dietary changes.<br>C) Agriculture Optimization<br>• Model Type: Predictive models (e.g., regression, CNNs for image-based plant disease detection), and prescriptive analytics for input usage.<br>• Goal: Optimize planting schedules, resource use, and harvest predictions.<br>All models will be fine-tuned using localized datasets and will leverage cloud-based platforms for scalability. |

# 2 Use Case Description

## 2.1 Description

The initiative addresses the challenges of food security, learner health, and personalized learning within higher educational institutions (HEIs) ecosystems. The primary objective is to integrate artificial intelligence into HEI ecosystems to improve food security, learner health, and personalized learning. The system combines real-time data collection (using IoT devices and wearables), AI-powered analytics, and adaptive learning platforms. This combination is used to optimize agricultural practices, improve nutrition, and deliver tailored educational content. The expected benefits include supporting the United Nations Sustainable Development Goals and promoting holistic learner development, mainly targeting HEIs in resource-constrained environments.

### Partners

- The Uganda National ICT Innovation Hub: Provision of workspace and internet connectivity, among others.
- The Uganda National Information Technology Authority (NITA), specifically the Personal Data Protection Office: Data collection and Data control approval.

## 2.2 Benefits of use case

- AI-driven financial inclusion strategies empower communities economically
- Smart agriculture and AI-driven nutritional insights enhance food security

- AI-powered diagnostics and personalized healthcare improve health outcomes
- Adaptive learning models provide customized educational experiences
- AI-driven job-matching

## 2.3 Future Work

The insights generated from data analysis lead to real-time and planned interventions in the following domains:

A) In Education:

- Personalized learning paths are recommended to learners based on knowledge gaps.
- Alerts educators when learners are disengaged or falling behind.
- Automated feedback and progress dashboards for learners and institutions.

B) In Health & Nutrition:

- Diet recommendations based on nutrient deficiencies detected.
- Health alerts to school/institutional medical staff if risk thresholds are detected.
- Meal program adjustments based on aggregated learner nutrition data.

C) In Agriculture:

- Automated irrigation or planting schedules based on weather and soil analytics.
- Advisories on pest control or fertilizer use.
- Reports to meal program managers linking crop readiness with school nutrition planning.

AI-driven personalization presents a powerful tool for advancing sustainable development across education, wealth, nutrition, and health sectors. The proposed framework enhances human well-being by integrating machine learning, IoT, and blockchain while ensuring economic and social equity. Future research will focus on real-world deployment, impact assessments, and scaling strategies for global implementation.

# 3 Use Case Requirements

**REQ-01** - Technical Requirements:

- AI Algorithms: Machine learning models for predictive analytics in agriculture, health monitoring, and adaptive learning.
- Data Integration: A secure, centralized platform for collecting and analyzing agricultural, nutritional, and educational data.
- Connectivity: Reliable internet access or offline-capable solutions for rural or underserved HEIs.
- Cloud Infrastructure: Scalable storage and computing power for real-time data processing.

**REQ-02** - Functional Requirements:

Wearables and Sensors Overview:

We are adopting commercially available and locally adaptable smart wearables and environmental sensors with the following capabilities:

A) Health & Nutrition Wearables (Student-focused):

- Capabilities: Heart rate monitoring, temperature, sleep patterns, hydration level estimations, step count, and real-time physical activity tracking.
- Examples: Devices similar to Fitbit, Huawei Band, or affordable alternatives that support health data APIs.
- Calibration: These are factory-calibrated but can be adjusted based on demographic norms (e.g., age or body mass index). Periodic validation with institutional health practitioners is planned.

B) Agriculture & Environmental Sensors (Meal Program Optimization):

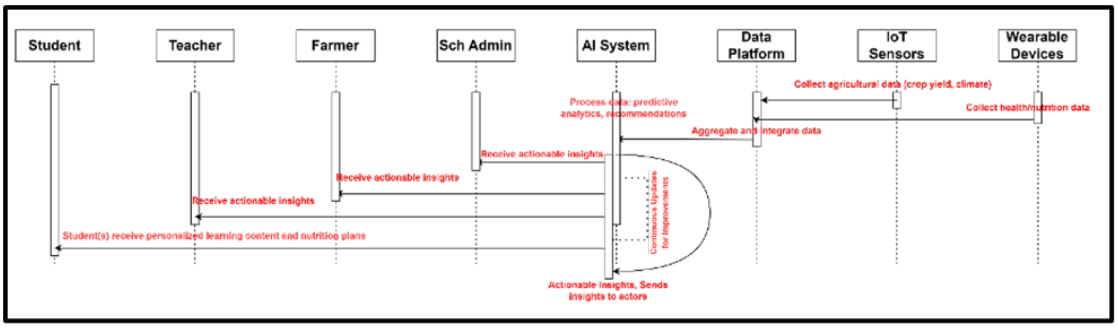
- Capabilities: Soil pH, temperature, humidity, light intensity, rainfall detection, and pest activity.
- Examples: Arduino-based and off-the-shelf IoT sensor kits.
- Calibration: Calibration follows manufacturer guidelines and is checked bimonthly, especially for field sensors, using standard reference measures and physical sample testing.

Data from these devices is integrated via local hubs (e.g., Raspberry Pi or mobile phones) and transferred to a cloud-based analytics system.

**REQ-03** - Operational Requirements:

- Training for staff in ICT department(s) and/or system administrators, on using the system.
- Collaboration with stakeholders, including government agencies, NGOs, and local communities.
- Policy frameworks for data privacy, ethical AI, and equitable access to the platform.
- Funding for pilot implementation in selected institutions.
- Long-term investment in scaling the platform across different regions.

## 4 Sequence Diagram




## 5 References

[1] AI and Agriculture: Wolfert, S., et al. (2017). Big Data in Smart Farming – A review. Agricultural Systems, 153, 69-80. DOI: 10.1016/j.agsy.2017.01.023

[2] AI in Education: Holmes, W., et al. (2019). Artificial Intelligence in Education: Promises and Implications for Teaching and Learning. OECD Publishing.

[3] https://doi.org/10.1787/9789264364070-en

[4] AI in Health and Nutrition: Krittanawong, C., et al. (2020). Machine Learning and Artificial Intelligence in Cardiovascular Medicine. Journal of the American College of Cardiology, 71(23),

[5] 2668-2679. DOI: 10.1016/j.jacc.2018.03.521

[6] Ethical AI: Jobin, A., et al. (2019). The Global Landscape of AI Ethics Guidelines. Nature Machine Intelligence, 1, 389-399. DOI: 10.1038/s42256-019-0088-2

# Use Case 11: Dubai AI & KHDA Smart School Finder

**Country:** United Arab Emirates

**Organization:** Dubai Digital Authority, Dubai Digital Government Establishment

**Contact Person**:

Hadi Ajeen (c_hadi.ajeen@digitaldubai.ae)

## 1 Use Case Summary Table

| Category | Education |
|---|---|
| The problem to be addressed | Users struggle to find suitable schools in Dubai due to fragmented data across different platforms and difficulty filtering based on personal criteria. |
| Key aspects of the solution | – Education Technology (EdTech), Government Services, AI Chatbot<br>– AI-powered chatbot interface for natural language school search<br>– KHDA API integration for real-time school data<br>– Personalized filtering based on user inputs (e.g., grade, fees, location, rating) |
| Technology keywords | Chatbot, NLP (Natural Language Processing), API Integration, School Search, KHDA, Education Analytics, AI, User Personalization |
| Data availability | Available via KHDA [2] APIs (schools list, locations, fees, inspection ratings, curriculum, contact details) |
| Metadata (type of data) | – School names<br>– Ratings<br>– Fees<br>– Curriculum types<br>– Grade levels offered<br>– Location coordinates |
| Model training and fine-tuning | – Pre-trained NLP model fine-tuned on school-related queries and KHDA taxonomy<br>– Intent classification for user queries |
| Testbeds or pilot deployments | Initial pilot within KHDA portal, Dubai Now app & Government Portals [1] |

## 2 Use Case Description

### 2.1 Description

The Dubai AI & KHDA Smart School Finder is an AI-powered platform designed to address the challenges parents face when searching for schools in Dubai. The current school search experience is fragmented and relies on static, outdated information that lacks personalization.

Parents often struggle to find reliable, real-time data about schools, and existing platforms do not offer smart filtering options to meet their preferences. Additionally, KHDA receives thousands of school-related inquiries, creating a need for a more efficient solution.

This AI-powered tool provides a real-time, data-driven search experience by integrating KHDA-approved data and offering personalized recommendations based on factors like location, curriculum, fees, and ratings. Users can simply interact with the system via natural language, asking specific queries such as, "Show me British curriculum schools in Al Barsha with high ratings." The system immediately returns accurate, up-to-date school details, saving time and providing transparency.

The AI solution offers several advantages over traditional systems, including faster access to verified information, smarter recommendations, and the ability to personalize the search. It simplifies the process for parents while reducing KHDA's manual intervention in handling inquiries. However, while the AI-based approach ensures more personalized and accurate results, it is still evolving in terms of addressing complex, niche queries.

**Use case status:** This use case is part of a larger product in development.

**Partners:** KHDA (Knowledge and Human Development Authority) [2]

## 2.2 Benefits of use case

The Dubai AI & KHDA Smart School Finder enhances access to accurate, transparent, and real-time information about schools, enabling parents to make informed decisions about their children's education. By leveraging AI, the tool ensures fair access to education opportunities, supporting Dubai's vision for a world-class education system.

The platform utilizes advanced AI and data-driven technologies to modernize the school search process, integrating real-time data analytics and intelligent recommendations.

It contributes to Dubai's digital transformation agenda by improving access to educational services through innovative infrastructure.

By providing equitable access to school information, the platform ensures that all families, regardless of socio-economic background, can find the best education options for their children. It reduces information gaps and ensures that educational opportunities are available to everyone, including underserved communities.

The tool promotes better urban planning by helping families find schools closer to their residences, reducing unnecessary travel and supporting sustainable urban development.

It enhances smart city initiatives by integrating AI-driven education services into Dubai's digital governance ecosystem.

The platform is developed through collaboration between Dubai Digital Authority, KHDA, and other stakeholders, demonstrating the power of public-private partnerships.

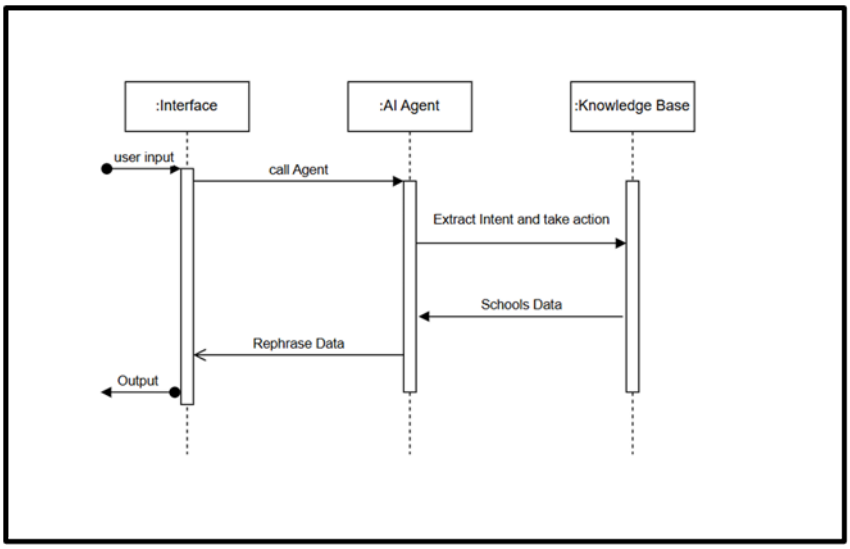
Future expansions may include partnerships with educational institutions, technology providers, and government entities to enhance data accuracy and service capabilities\

## 2.3 Future Work

- Data Collection: Expanding the database to include more granular school details such as extracurricular activities, special programs, and detailed fee structures.
- Proof of Concept Development: Further enhancing the platform's capabilities with additional AI-driven features, such as predictive analytics to forecast demand trends and admission chances.
- Model Development: Refining machine learning models to better understand user preferences and predict the best schools based on past interactions.
- Create New Variations/Extensions: Develop mobile app versions and multi-language support for wider adoption.
- Standards Development: Creating educational data standards for seamless integration with other government portals and educational platforms.

The future work would aim to improve data collection, enhance AI functionalities, and expand the platform to serve a broader user base, ensuring the tool becomes an integral part of the educational experience for all Dubai residents.

# 4 Sequence Diagram




# 5 References

[1] https://chat.dubai.ae/ - Test deployment of the chatbot

[2] https://web.khda.gov.ae/en/ - KHDA Website for resources and further information

# 4.8. Category 8: Smart home/cities

## Use Case 1: Culturalink: AI-Driven World Cultural Linking and Interactions Engine

**Organization**: Long-term AI

**Country**: China

**Contact Person(s)**:

Professor Yi Zeng, yi.zeng@ia.ac.cn

Kang Sun, kang.sun@long-term-ai.cn

Qian Liang, qian.liang@ia.ac.cn

### 1 Use Case Summary Table

| Item | Details |
|---|---|
| Category | Smart home/cities |
| Problem Addressed | The primary challenge is enabling AI to accurately analyze, link, and interpret vast, diverse cultural heritage data across regions and formats while ensuring cultural sensitivity, inclusiveness, and global interoperability. |
| Key Aspects of Solution | The proposed solution integrates a dual-engine architecture—comprising a global cultural data engine and an AI-driven intelligence engine—to perform multimodal analysis, cultural classification, cross-heritage linking, creative content generation, and dynamic cultural knowledge mapping. |
| Technology Keywords | Cultural analysis, Cultural interaction, AI-powered culture heritage, Multimodal cultural data process, Cultural Knowledge Graph |
| Data Availability | Private |
| Metadata (Type of Data) | Texts, Images |

(continued)

| Item | Details |
|---|---|
| Model Training and Fine-Tuning | The platform leverages AI/ML technologies for intelligent cultural data classification, automated cross-heritage correlation analysis, and creative content generation. Its architecture integrates a large-scale cultural knowledge base with machine learning–driven semantic mapping and dynamic cultural evolution tracking. |
| Testbeds or Pilot Deployments | https://culturalink.ai/ |
| Code repositories | None |

# 2 Use Case Description

## 2.1 Description

Culturalink[1] is an AI-powered global cultural analysis, linking, and interaction system platform designed to deeply explore and connect material cultural heritage, intangible cultural heritage, and other cultural content worldwide. Through multimodal data analysis and knowledge graph construction technologies, the platform reveals hidden connections across regions, cultures, and eras, showcasing the profound essence of cultural integration. It leverages AI to assist in the preservation and inheritance of cultural heritage, empower cultural innovation, and foster dialogue between tradition and modernity.

With a vision of intertwining human and artificial intelligence to weave a global cultural network that connects history to create the future, Culturalink operates under the core principles of fairness and inclusiveness. It provides equal dissemination and protection opportunities for all cultural groups worldwide, promoting harmonious coexistence of diverse cultures by identifying commonalities and connections among different cultural expressions. The platform contributes to the diversity and sustainable development of human civilization.

The Platform Architecture of Culturalink include both Global Cultural Data Engine and AI-driven Cultural Intelligence Engine. The Global Cultural Data Engine is responsible for the structured ingestion, semantic integration, and alignment of diverse cultural heritage data collected from over 190 countries. It processes multimodal inputs, primarily textual and visual data, from the following four major sources: (1)Web-crawled open data that includes both text (e.g., encyclopedic descriptions, forum discussions, web articles) and image data (e.g., artifact photos, traditional motifs) harvested from publicly accessible online sources such as websites, forums, and encyclopedias; (2)Internally curated knowledge bases, such as domain-specific textual descriptions (e.g., structured metadata, historical records) and corresponding visual materials (e.g., high-resolution scanned images, exhibition photographs) that have been compiled through prior research and institutional archival efforts; (3)Authoritative open dataset contains texts and images from internationally recognized sources such as UNESCO World Heritage listings, Hugging Face-hosted cultural corpora, and Wikidata; (4)User-contributed data submitted through the platform interface, including textual narratives, annotations, and associated visual content from users(e.g., community-documented images and folk artifact photos). The AI-driven Cultural Intelligence Engine supports intelligent analysis, narrative generation, and user-facing interaction by integrating multiple capabilities into a unified framework. It performs multimodal analysis[3] by cross-referencing text and image data to

extract semantic entities, relationships, and narrative dimensions; for example, a silk artifact can be semantically linked to trade routes (spatial), production techniques (processual), and mythological symbolism (textual). Through cross-heritage linking, the system applies graph embeddings, similarity computations, and large language model (DeepSeek)-based reasoning to uncover implicit relationships between disparate cultural entities—for instance, associating a ritual dance with textile patterns through shared symbolic motifs. In terms of creative content generation, the engine employs LLMs with structured prompt design to produce explanatory text, narrative paragraphs, and multilingual summaries, while text-to-image models are used to reconstruct relevant scenes, tools, or environments and image style transfer techniques grounded in cultural context. This integrated architecture enables dynamic thematic exploration—for example, constructing a multi-dimensional narrative around "Beijing Central Axis," which interweaves historical events, regional cultural practices, and material heritage—facilitated through AI-assisted storytelling and interactive presentation mechanisms.

**Use Case Status:** Part of a larger research project and larger product development

**Partners:** Brain-inspired Cognitive AI Lab, at Institute of Automation, Chinese Academy of Sciences (https://braincog.ai/)

## 2.2 Benefits of use case

Culturalink promotes global cultural literacy by providing AI-driven insights into tangible and intangible heritage and their associations, from humanity, culture, literature, and philosophy perspectives. It educates users about diverse cultures, fostering empathy, and advancing education for global citizenship. It helps stakeholders understand interconnected cultural narratives.

- By harnessing frontier AI models, Culturalink demonstrates how AI can be applied to preserve and analyze cultural data, mapping connections between disparate cultural practices, setting a precedent for tech-driven heritage conservation.
- By highlighting underrepresented cultures and emphasizing interconnectedness, Culturalink reduces cultural marginalization. It challenges stereotypes and promotes equity in cultural representation. Amplifying intangible heritage and ensuring diverse voices are recognized.
- Directly protects cultural/heritage sites. Culturalink digitally preserves endangered heritage and raises awareness about their significance. Virtual linkages between heritage sites promote sustainable tourism and community pride.
- Fosters intercultural dialogue and mutual respect showcasing shared human values, it strengthens social cohesion. Philosophical analyses reveal universal themes across cultures, bridging divides.

## 2.3 Future Work

Culturalink will continue to reshape its data and knowledge engine, strengthen its AI engine for domestic and international cultural relics, historical archives, regional characteristics, and philosophical thought systems. Leveraging AI, it will deeply explore the spatiotemporal correlations among cultural artifacts and heritage sites worldwide, conduct cross-cultural interpretations of philosophical concepts, and analyze the spatiotemporal connections between tangible and intangible cultural heritage to enhance understanding of regional civilization

dissemination and evolution. This will reveal the chronological and spatial trajectories of cultural development.

The development roadmap and corresponding timeline are outlined as follows:

Short-term (within 6–12 months):

The platform will deepen its data acquisition efforts across more countries and cultural contexts to ensure data diversity, authenticity, and high quality. This expanded coverage is essential for reducing cultural bias, enhancing fair representation, and enabling more inclusive AI-generated narratives. Meanwhile, the AI engine will be optimized for region-aware narrative generation and multilingual support, particularly for non-Western and low-resource languages. The semantic alignment pipeline will also be upgraded to strengthen cross-domain linking between tangible and intangible cultural elements, supporting a more comprehensive understanding of global cultural interconnections.

Mid-term (1–2 years):

The platform will focus on the automated generation of cultural narratives centered around thematic topics. By leveraging AI-driven semantic analysis and narrative synthesis, it will support the construction of cross-cultural storylines grounded in shared motifs, historical trajectories, and regional philosophies. A user-facing interface will allow dynamic generation of these thematic narratives, enabling personalized cultural exploration. In parallel, visual generation models will be enhanced with culturally adaptive style transfer capabilities to ensure that reconstructed scenes, tools, and artifacts align with their respective cultural contexts, thereby improving the expressiveness and authenticity of multimodal storytelling.

Long-term (2+ years):

The system will evolve into an open, modular infrastructure that facilitates collaboration across global cultural research communities. It will offer interfaces for cultural scholars, educators, and content creators to contribute, query, and reinterpret cultural data across regions. To ensure high-impact deployment and governance, the platform aims to establish strategic partnerships with national cultural institutions and international organizations—such as UNESCO, ICOM, or regional heritage networks—to jointly promote cultural data standardization, knowledge sharing, and ethical AI use in heritage preservation. Deep learning pipelines will be refined to model the evolution of cultural elements over time and space, revealing patterns of convergence, divergence, and fusion among civilizations, while ensuring alignment with globally recognized cultural diversity and sustainability goals.

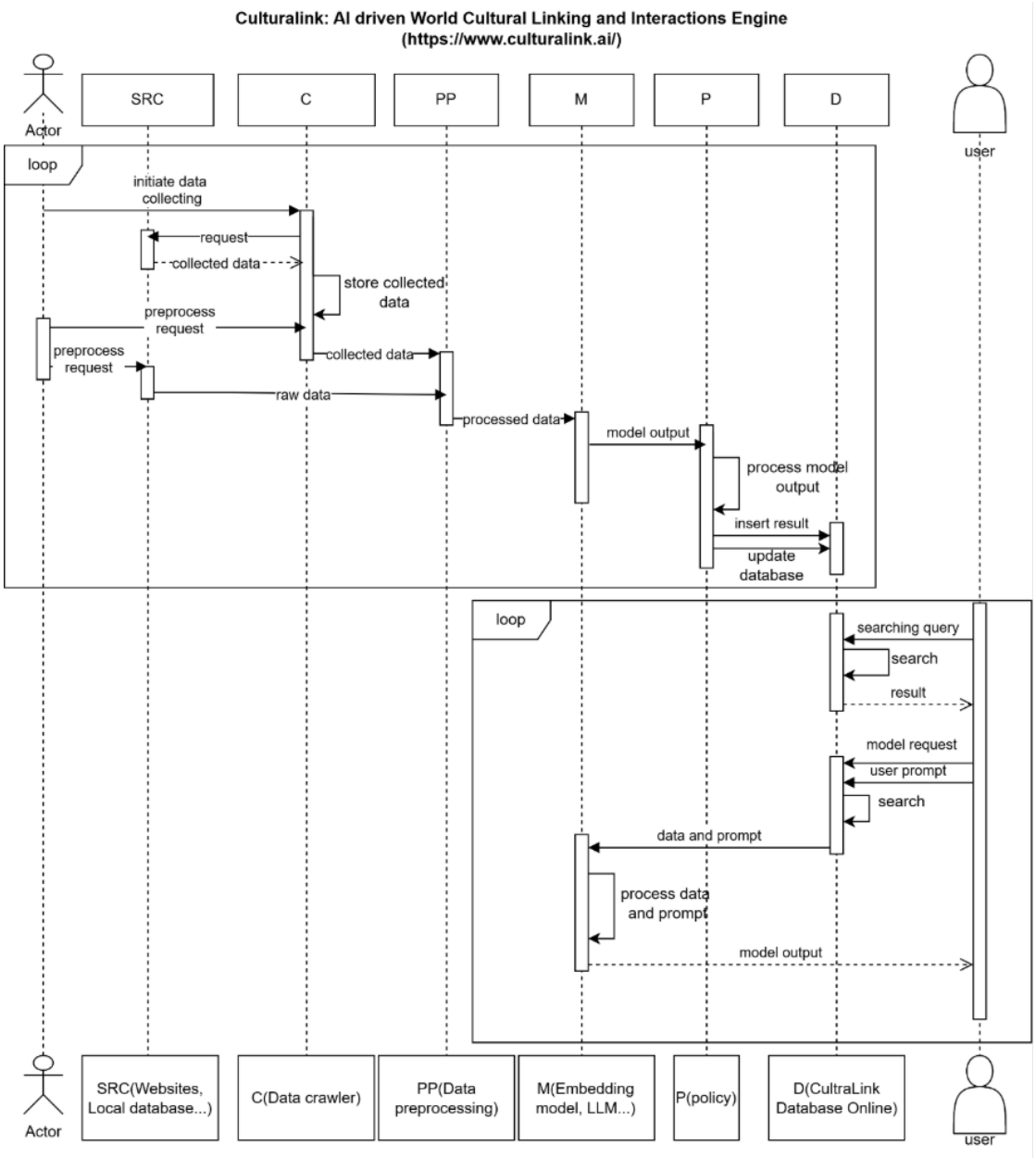
The Culturalink platform particularly emphasizes the preservation and transmission of intangible cultural heritage, supporting the continuity and innovation of cultural diversity. It serves not only as an AI-powered cultural exploration tool but also as a cultural bridge connecting past, present, and future. Through multidimensional analysis and digital archiving technologies, Culturalink fosters inclusive and sustainable development of global cultures, enabling users to trace cultural diffusion paths across civilizations while providing innovative solutions for heritage conservation and cultural industry development.

## 3 Use Case Requirements

- **REQ-01**: The system must support the ingestion and integration of large-scale, multimodal cultural data from at least 172 countries, encompassing both tangible and intangible heritage sources. This includes structured metadata (e.g., from national heritage databases), unstructured narratives (e.g., oral histories and folklore), and visual content (e.g., scanned manuscripts, images of museum artifacts). To ensure inclusivity and linguistic diversity, the system will initially provide multilingual support covering at least five low-resource languages.
- **REQ-02**: The platform will integrate advanced techniques in natural language processing, computer vision, audio signal processing, and video analysis to enable comprehensive interpretation of multimodal cultural content. For NLP tasks such as named entity recognition and thematic modelling, models like DeepSeek and a proprietary fully spiking brain-inspired language model[2][7] will be used to capture fine-grained cultural semantics. In the visual domain, image classification and style transfer may leverage models such as CLIP, Segment Anything Model (SAM), or Stable Diffusion for cross-modal grounding and aesthetic reconstruction. Audio processing will employ techniques such as rhythm and timbre extraction, with speech recognition supported by models like Whisper. Video understanding will incorporate keyframe detection and scene segmentation using vision transformers or similar architectures[5]. Together, these components will support automated metadata extraction—including object names, cultural symbols, spatial references, and historical contexts—thereby enabling a cross-modal semantic labeling framework that drives downstream cultural knowledge modelling and AI-assisted narrative generation[6].
- **REQ-03**: A temporospatial knowledge neural network will be constructed to represent cultural entities (e.g., artifacts, rituals, people, places) and their relationships across historical periods and geographic regions. The system will utilize graph databases (e.g., Neo4j) for scalable representation and integrate relational graph convolutional networks (e.g., R-GCNs) for embedding and inference. This will enable advanced queries such as tracing the origin and evolution of a specific ceramic type or identifying the influences behind a regional folk festival.
- **REQ-04**: The platform will implement large language models (e.g., BERT or GLM) combined with cultural knowledge neural network to generate semantic representations of cultural artifacts and content. This enables automatic categorization into thematic groups such as “agricultural heritage objects,” “Silk Road-related traditions,”and etc.
- **REQ-05**: Using graph embeddings, similarity metrics, and reasoning capabilities of LLMs, the system will uncover implicit relationships between disparate cultural elements. For instance, it may link a traditional dance with a textile motif based on shared symbolic meaning. Semantic correlation will be guided through prompt-based inference and knowledge graph traversal, supporting cross-domain and cross-cultural exploration of heritage elements.
- **REQ-06**: The platform will use large language models (e.g., our proprietary fully spiking brain-inspired language model or DeepSeek) with structured prompts and culture knowledge neural network to generate explanatory texts, narrative segments, and multilingual summaries for cultural interpretation. Text-to-image models (e.g., Stable Diffusion) will be used to reconstruct historical scenes[4], production techniques, and environmental settings. Additionally, style transfer techniques will simulate different artistic styles across time and geography, enhancing the visual richness of AI-generated content in exhibitions or digital storytelling.
- **REQ-07**: The system will support temporospatial analysis of cultural phenomena through temporal modelling (temporal dimension of culture knowledge neural network) and spatial representation (e.g., space dimension of culture knowledge neural network). It will enable visualization and inference of cultural transitions—such as how a particular pottery

style spread across regions or how a ritual evolved across centuries—providing valuable insights for research in cultural transmission, divergence, and continuity.

## 4 Sequence Diagram




## 5 References

[1] Culturalink. Accessed August 25, 2025. https://culturalink.ai/en.

[2] Zeng, Yi, Dongcheng Zhao, Feifei Zhao, Guobin Shen, Yiting Dong, Enmeng Lu, Qian Zhang, Yinqian Sun, Qian Liang, Yuxuan Zhao, Zhuoya Zhao, Hongjian Fang, Yuwei Wang, Yang Li, Xin Liu, Chengcheng Du, Qingqun Kong, Zizhe Ruan, and Weida Bi. 2023. "BrainCog: A Spiking Neural Network Based Brain-inspired Cognitive Intelligence Engine for Brain-inspired AI and Brain Simulation." Patterns. Cell Press.

[3] He, Xiang, Xiangxi Liu, Yang Li, Dongcheng Zhao, Guobin Shen, Qingqun Kong, Xin Yang, and Yi Zeng. 2024. "CACE-Net: Co-guidance Attention and Contrastive Enhancement for Effective Audio-Visual Event Localization." ACM Multimedia.

[4] Shen, Guobin, Dongcheng Zhao, Xiang He, Linghao Feng, Yiting Dong, Jihang Wang, Qian Zhang, and Yi Zeng. 2024. "Neuro-Vision to Language: Image Reconstruction and Interaction via Non-invasive Brain Recordings." Advances in Neural Information Processing Systems (NeurIPS).

[5] He, Xiang, Xiangxi Liu, Yang Li, Dongcheng Zhao, Guobin Shen, Qingqun Kong, Xin Yang, and Yi Zeng. 2024. "CACE-Net: Co-guidance Attention and Contrastive Enhancement for Effective Audio-Visual Event Localization." ACM Multimedia.

[6] He, Xiang, Dongcheng Zhao, Yang Li, Guobin Shen, Qingqun Kong, and Yi Zeng. 2024. "An Efficient Knowledge Transfer Strategy for Spiking Neural Networks from Static to Event Domain." Proceedings of the AAAI Conference on Artificial Intelligence.

[7] Shen, Guobin, Dongcheng Zhao, Yiting Dong, Yang Li, Jindong Li, Kang Sun, and Yi Zeng. 2023. "Astrocyte-Enabled Advancements in Spiking Neural Networks for Large Language Modeling." arXiv preprint arXiv:2023.

# Use Case 2: Yuanjing Integrated Urban Management Platform

**Organization**: China Unicom

**Country**: China

**Contact Person(s)**:

Name: Liu Zhicen Email Address: liuzc158@chinaunicom.cn

Name: Sang Haiyan Email Address: sanghy6@chinaunicom.cn

Name: Guan Xiaoran Email Address: guanxr6@chinaunicom.cn

## 1 Use Case Summary Table

| Item | Details |
|---|---|
| Category | Smart home/cities |
| Problem Addressed | Clearly describe the primary issue or challenge<br>1.Low Governance Efficiency<br>Difficulties in cross-domain coordination: Urban management involves multiple departments such as transportation, environmental protection, and public safety. Traditional manual processing methods result in inefficient interdepartmental collaboration and delayed information transmission.<br>Slow response speed: Reliance on manual inspections and paper-based workflows leads to prolonged problem identification and resolution cycles.<br>2.Inadequate Public Service Quality<br>Delayed handling of citizen requests: Traditional hotline and counter services are inefficient, resulting in extended processing times for complaints and suggestions, leading to low satisfaction rates.<br>3.Excessive Operational Costs<br>High management costs due to substantial manpower investment and repetitive tasks. |

(continued)

| Item | Details |
|---|---|
| Key Aspects of Solution | List main elements of the proposed solution, including the role of AI.<br>1.Large Language Models (LLMs):<br>Handle multilingual citizen inquiries with 24/7 responsiveness; automatically parse complaint tickets to extract key information (location/issue type) and generate resolution recommendations.<br>2.Multimodal Large Models:<br>Simultaneously analyze unstructured data uploaded by citizens (e.g., "text descriptions + on-site photos") to accurately identify complex, hybrid issues.<br>3.Visual Large Models:<br>Monitor urban issues in real-time via cameras and promptly provide early warning alerts.<br>4.RAG (Retrieval-Augmented Generation):<br>Dynamically retrieve the latest policy documents to ensure AI responses remain compliant; match historical similar cases to provide actionable resolution templates. |
| Technology Keywords | Relevant technical keywords<br>LLMs, Multimodal large model, RAG, Fine-tuning |
| Data Availability | Public / Private (specify sources or repositories if available)<br>Private - Urban data |
| Metadata (Type of Data) | Specify data types (e.g., text, audio, visual)<br>Structured and unstructured data including natural language text, policies, reports, video and image |
| Model Training and Fine-Tuning | Briefly describe AI/ML technologies and approaches used 1.Pretrain with Urban data<br>2.Supervised Fine-Tuning(SFT) |
| Testbeds or Pilot Deployments | Provide links or references to deployment examples<br>It is private. |
| Code repositories | Provide github links (if available)<br>It is private. |

# 2 Use Case Description

## 2.1 Description

Yuanjing Integrated Urban Management Platform is based on advanced large-model technology, aiming to promote sustainable development in cities worldwide, enhance urban governance efficiency, optimize public service responsiveness, and create a smarter and more efficient living and working environment for citizens. By leveraging language large models, vision large models, and multimodal large models, the platform has implemented functions such as intelligent Q&A for public services, rapid dispatch of requests, precise detection of urban events, and real-time early warning and scheduling. With its efficient multi-mode data processing capabilities, the platform also supports the expansion of large-scale application

scenarios. These capabilities address the traditional challenges of slow information processing and low resource allocation efficiency in urban management, resulting in a 25% improvement in urban management efficiency in practical applications and provide scientific decision-making support for grassroots governance.

The platform fully considers international requirements, supporting multilingual interaction and cross-cultural adaptability. It also supports RAG to use different city background data. Every city can establish its own knowledge base,and it allows for customized deployment according to the urban governance needs of different countries and regions. The platform integrates external knowledge bases with large language models (LLMs) through RAG, establishing a dynamically collaborative data ecosystem. Platform administrators can store frequently asked questions, standardized answers, and relevant policies in the knowledge base in both structured and vectorized formats. Policy texts and Q&A pairs are transformed into high-dimensional vectors via embedding models and stored in a vector database for semantic retrieval, while other structured data is maintained in a relational database. For example, when a user submits a query, the platform first retrieves matching policy excerpts and case studies through vector similarity search, then performs logical validation using structured data, ultimately generating accurate responses and triggering corresponding actions such as automatic work order dispatch. User feedback and expert review results continuously optimize the knowledge base ranking and model parameters, ensuring the system maintains high accuracy even after policy updates.h

Due to data privacy concerns, the platform can be privately deployed on government-owned servers. Alternatively, we offer connectivity to China Unicom cloud servers, which provide high-level security and trusted data circulation technologies. To ensure the credibility of AI reasoning and the effectiveness of actions, the platform adopts a closed-loop system combining "expert oversight + transparency mechanisms + tiered response": First, domain experts manually review high-risk decisions, and models are continuously refined through feedback. Additionally, all AI outputs must include explainable justifications (e.g., policy clauses, similar cases). The platform automatically implements tiered responses based on reasoning confidence: high-confidence inferences trigger immediate execution, while low-confidence cases are escalated for human intervention or require supplementary evidence.

Although the initial deployment of the platform may face issues such as high hardware costs, continuous optimization of algorithms and model scheduling can gradually reduce resource consumption and expand the scope of application. Meanwhile, with the advancement of communication technologies and the widespread adoption of 5G/6G, the platform will achieve lower-latency responses and support the integration of more urban devices.

**Use Case Status:**

The use case is part of a larger product development.

## 2.2 Benefits of use case

The Yuanjing Integrated Urban Management Platform, through the application of advanced large model technology, enhances urban management efficiency and optimizes resource allocation, providing long-term value for the sustainable development goals of cities and communities.

For example, the intelligent dispatch and event closed-loop management functions can quickly respond to citizens' requests, promote the equitable accessibility of basic services, and pay particular attention to the needs of vulnerable groups, thereby fostering inclusive urbanization.

The platform's real-time event detection and intelligent early warning functions help improve social governance efficiency, reduce resource waste, and lower labor demand and operational costs through intelligent tools, thus driving green urban development. Moreover, with its efficient data processing capabilities based on large models and the ability to expand into large-scale application scenarios, the platform is more universally applicable and suitable for promotion in different urban environments around the world.

### 2.3 Future Work

Work Plan for the First Half of 2025:

Model Refinement and Miniaturization: Develop specialized small models tailored to the specific needs of urban and community management, ensuring efficient operation in resource-constrained environments.

Multimodal Data Integration: Strengthen the platform's ability to process multimodal data (text, images, videos), enhancing understanding and analysis of complex urban scenarios, and providing more comprehensive intelligent services.

Work Plan for the Second Half of 2025:

Scene-driven Model Optimization: Continuously optimize model performance based on actual application scenarios and user needs, iterating the product's capabilities to ensure effective implementation in urban and community management, and expanding its reach to more regions.

Open Source Sharing and Ecosystem Building: Actively participate in the open-source community, share technical achievements, build a collaborative ecosystem, and promote the widespread application of large model technology in sustainable urban and community development.

## 3 Use Case Requirements

**REQ-01**: Improving Urban Governance Efficiency

Urban management involves multiple domains such as transportation, environmental protection, and public safety. Traditional manual processing methods are inefficient, urgently requiring intelligent solutions to achieve automated task dispatching, intelligent analysis, and rapid response, thereby reducing administrative costs and enhancing problem-solving speed.

**REQ-02**: Optimizing Public Services

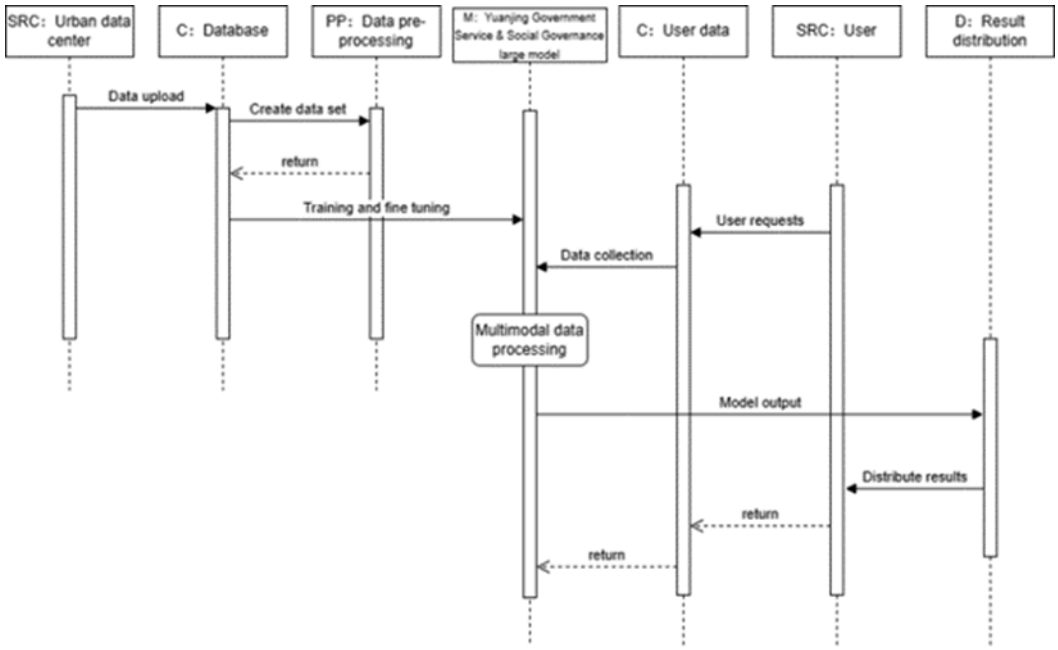
Citizens' expectations for convenient and efficient public services are growing. Cities need to leverage intelligent customer service, automated task dispatching, mobile feedback functions, and other smart solutions to shorten request processing time and improve citizen satisfaction.

**REQ-03**: Reducing Operational Costs

Traditional urban management relies heavily on manual inspections and redundant construction, resulting in high costs. Cities require intelligent solutions to optimize resource allocation (e.g., garbage truck route planning), minimize waste, and enhance the efficiency of fiscal fund utilization.

## 4 Sequence Diagram

- Insert a sequence diagram or conceptual diagram clearly illustrating interactions between different components or users within the use case.




## 5 Reference

[1] United Nations (UN). Available online: https://china.un.org/en/sdgs/11

# Use Case 3: AI-Enhanced Detection and Prediction at Saruq Al-Hadid Archaeological Site, Dubai, UAE

**Organization:** Dubai Culture & Arts Authority, Khalifa University of Science and Technology

**Country:** United Arab Emirates

**Contact Person(s):**

Primary: Mariam Ali Al Suwaidi, Mariam.Alsuwaidi@dubaiculture.ae

Secondary: Diana Francis, diana.francis@ku.ac.ae

Additional Contacts: Bader Mohamed Al Ali, Bader.Alali@dubaiculture.ae

Additional Contacts: Luma Bourisly, Luma.bourisly@dubaiculture.ae

## 1 Use Case Summary Table

| Item | Details |
|---|---|
| Category | Smart home/cities |
| Problem Addressed | The primary issue addressed is the challenge of accurately detecting and predicting buried archaeological sites in arid, desert environments. Traditional ground-based methods often struggle in these areas due to extensive sand cover, complex stratigraphy, and limited surface visibility. These factors hinder efficient site discovery, delay excavation planning, and increase operational costs, especially in environmentally sensitive or logistically complex terrains such as the Saruq Al-Hadid site in Dubai, United Arab Emirates (UAE). |

(continued)

| Item | Details |
|---|---|
| Key Aspects of Solution | The approach tackles these challenges by integrating advanced remote sensing data (from Synthetic Aperture Radar (SAR) and high-resolution multispectral imagery) with machine learning and geospatial analysis techniques. The goal is to automate the detection process and improve guidance for future archaeological investigations at sites like Saruq Al-Hadid in Dubai, UAE.<br>Key elements of the proposed solution include:<br>Multi-Sensor Remote Sensing Integration: Utilization of high-resolution multispectral imagery (Worldview-3) and SAR data (Advanced Land Observing Satellite (ALOS)-2/Phased Array type L-band Synthetic Aperture Radar (PALSAR)-2) to enhance surface and subsurface feature detection.<br>Advanced Image Processing and Feature Enhancement: Application of Principal Component Analysis (PCA), spectral band transformations, and band ratio indices (e.g., Normalized Difference Vegetation Index (NDVI), Normalized high-frequency difference (NHFD)index to enhance and extract meaningful features from the remote sensing data.<br>Application of artificial intelligence: Implementation of machine and deep learning algorithms—including unsupervised classification (e.g., Iterative Self-Organizing Data Analysis Technique Algorithm (ISODATA)), K-means++ clustering) and convolutional neural networks (CNNs) for semantic segmentation to automatically detect features like dune patterns and potential archaeological features.<br>Geospatial Analysis and Data Integration: Extraction and processing of geospatial data (e.g., digital elevation models, slope maps, hydrographic networks and dune morphology) combined with geostatistical methods to analyze spatial patterns and relationships among various landscape elements.<br>Multimodal Reclassification and Expert Input: Fusion of remote sensing data, geospatial analysis, and field expertise to reclassify the landscape into distinct geomorphological assemblages that distinguish between natural formations and potential archaeological sites.<br>Predictive Modeling and Pattern Recognition: Development of predictive models using techniques such as Kernel PCA and Gaussian radial basis function (RBF) kernels to identify and prioritize areas of potential archaeological significance, guiding future on-site investigations.<br>Validation through field data, excavation reports, and supporting geophysical surveys—including Ground Penetrating Radar (GPR) and magnetic survey—helped confirm the accuracy of the predictions.<br>These components collectively facilitate more efficient and automated archaeological prospection. |
| Technology Keywords | Remote Sensing, Synthetic Aperture Radar (SAR), Multispectral Imagery (Worldview-3, ALOS-2/PALSAR-2), Spectral Band Transformation, Band Ratio Indices (NDVI, NHFD), Principal Component Analysis (PCA), Advanced Image Processing, Unsupervised Classification (ISODATA, K-means++), Machine Learning, Deep Learning, Convolutional Neural Networks (CNNs), Gaussian RBF Kernels, Geospatial Analysis, Digital Elevation Models (DEMs), Geographic Information Systems (GIS), Archaeological Site Detection, Pattern Recognition, Data Integration, Data Fusion, Ground Penetrating Radar (GPR), Magnetic Survey, Geophysical Validation. |
| Data Availability | Public |

(continued)

| Item | Details |
|---|---|
| Metadata (Type of Data) | Mix of data types, including:<br>Visual Data: High-resolution Red-Green-Blue (RGB) ortho-mosaics and classified index maps (e.g., NDVI, NHFD). Remote Sensing Data: High-resolution multispectral images (e.g., World-view-3), synthetic aperture radar (SAR) imagery from ALOS-2/PALSAR-2. Geospatial Data: GIS layers (elevation, slope, hydrology, land cover, excavation zones) and Global Positioning System (GPS)-based ground control points and topographic grids<br>Geophysical data: Ground Penetrating Radar (GPR) and magnetic survey results used for subsurface anomaly detection and validation of remote sensing predictions. Textual Metadata: Archaeological excavation reports, documen-tation of processing techniques and methodological details. Articles have been published in open-access journals. Further data sharing may be possible upon request. |
| Model Train-ing and Fine-Tuning | The use case employs several technologies and approaches for model training and fine-tuning, including:<br>Unsupervised Learning: Applied ISODATA and K-means++ clustering to group multispectral and SAR data based on spectral and spatial similarity without predefined labels. Used PCA and Kernel PCA for dimensionality reduction and to identify dominant geospatial patterns. Kernel PCA was combined with a Gaussian Radial Basis Function (RBF) kernel to enhance predictive modeling of archaeo-logical zones.<br>Deep Learning: implemented U-Net (CNN) for semantic segmentation of dune patterns from labeled satellite imagery. The model was trained on 120 RGB images using an 80/20 train-test split, with binary cross-entropy loss and binary accuracy as the evaluation metric. Training was stabilized over 30 epochs, achiev-ing a validation accuracy of 95%.<br>Processing Frameworks: Model development and training were conducted using the Keras framework on an NVIDIA A40 Graphics Processing Unit(GPU) within a high-performance computing environment.<br>Workflow integrated into ArcGIS Pro for spatial analysis, deployment, and result visualization.<br>Model refinement was supported by feedback from field validation, including results from ground-penetrating radar (GPR) and magnetic surveys, which helped adjust the model's predictive accuracy in real-world conditions.<br>These combined approaches allowed for the effective training and fine-tuning of models to accurately detect and predict archaeological features using remote sensing data. |

(continued)

| Item | Details |
|---|---|
| Testbeds or Pilot Deployments | The use case was implemented as a pilot deployment at the Saruq Al-Hadid site in Dubai, UAE. This site served as the testbed for evaluating the efficiency and accuracy of the integrated remote sensing, geospatial analysis, and machine learning framework. Field validation and comparisons with historical excavation data were used to assess the performance of the approach, demonstrating its potential to guide targeted archaeological investigations in challenging arid environments. These results validate the integrated AI–remote sensing framework in a real-world desert context. In addition, supporting geophysical surveys—including Ground Penetrating Radar (GPR) and magnetic survey—conducted in parallel for validation revealed further insights into the site's subsurface anomalies. Key findings from the AI–remote sensing and machine learning approach include the successful identification of known excavation areas and the detection of new locations with archaeological potential. Meanwhile, geophysical surveys at SAR.53 identified a total of 462 anomalies, including five major subsurface structures (A–E), while SAR.7 yielded 2,094 anomalies distributed across six sub-zones with archaeological potential (A–F).<br>Links: [1] |

# 2 Use Case Description

## 2.1 Description

Archaeological research in arid regions such as the Saruq Al-Hadid site in Dubai, UAE, is challenged by extensive sand cover, shifting dunes, and minimal vegetation, which limit the effectiveness of traditional field surveys in detecting buried features. This use case addresses these limitations through an innovative, integrated approach that combines high-resolution remote sensing technologies—specifically multispectral imagery and SAR—with advanced image processing, machine learning, and geospatial analysis. The objective is to automate the detection and prediction of archaeological sites, reducing reliance on invasive, labor-intensive surveys while guiding excavations more efficiently and accurately. The technical approach includes PCA, spectral band transformations, and semantic segmentation using convolutional neural networks (U-Net), alongside clustering techniques such as K-means++ and Kernel PCA with Gaussian RBF kernels. These tools extract meaningful patterns from complex datasets. Geospatial tools, including digital elevation models and slope maps, help contextualize findings within the broader landscape. The framework was further validated through comparisons with historical excavation data and supported by geophysical surveys—namely GPR and magnetic survey—offering deeper insights into subsurface features. The solution has demonstrated high detection accuracy, successfully identified known excavation areas and predicted new zones with archaeological potential. It enhances decision-making and resource allocation by helping archaeologists focus efforts on the most promising areas. The methodology is also scalable and adaptable for use in other arid or environmentally sensitive regions. Compared to traditional survey methods, it offers a major advancement by enabling automated, non-invasive, and data-driven archaeological investigations at both surface and subsurface levels.

The outputs of this approach contributed to the creation of a site-specific geospatial knowledge base for the Saruq Al-Hadid archaeological site. This includes digital elevation and slope maps, hydrographic network data, classified geomorphological assemblages (Geo Assemblage I–III), and AI-generated outputs such as unsupervised classifications, clustering results, and predicted

archaeological zones. These structured datasets, derived from both remote sensing analysis and field validation, support archaeological interpretation, spatial modeling, and decision-making—reinforcing the approach's value in both research and heritage management.

**Use Case Status:** Pilot completed. The use case has moved beyond the initial research phase and was successfully implemented and tested at the Saruq Al-Hadid archaeological site in Dubai, UAE, serving as a full-scale pilot to validate the proposed methodology.

**Partners:** Dubai Culture & Arts Authority, Khalifa University.

## 2.2 Benefits of the use case

The use case delivers significant and multifaceted benefits across education, research, conservation, and public sector innovation. It advances academic knowledge and supports training in archaeological and remote sensing techniques, contributing to quality education and lifelong learning. As a non-invasive innovation, it enabled the detection of archaeological anomalies at the Saruq Al-Hadid site without physical excavation—preserving site integrity and ensuring sustainable conservation practices. The approach reduced fieldwork time and cost, enhanced operational efficiency, and guided excavations using high-resolution maps and 3D digital models. Strategically, it aligns with national goals in cultural sustainability, scientific research, and smart urban development. The project also encouraged interdisciplinary collaboration among governmental authorities, academic institutions, and data providers, fostering effective partnerships and knowledge exchange. It strengthened the role of cultural authorities as public sector innovation leaders and introduced a framework for archaeology–AI collaboration. Marking a shift from traditional manual surveys to AI-driven techniques rarely adopted in the region, the initiative is the first in the world to fully integrate satellite remote sensing, machine learning, and geophysical data (GPR and magnetometry) for archaeological discovery. This service and policy innovation introduced AI-powered site detection as a new exploration method, creatively combining diverse datasets into a unified, predictive workflow. The methodology is scalable and adaptable, supporting replication by other emirates, cultural bodies, or global heritage institutions, and applicable across sectors such as environment, tourism, and smart city planning. It is ideal for heritage sites under urban infrastructure or in fragile ecosystems and provides a blueprint for AI-powered cultural resource management. The project enables efficient assessments without disrupting communities or delaying development, supports academic publishing and public engagement, and contributes to preserving archaeological heritage for future generations. Ultimately, it exemplifies how AI can enhance cultural preservation, marking not only a technological breakthrough but a milestone in global heritage conservation.

## 2.3 Future Work

Model Enhancements and Generalization:

Future improvements will focus on integrating advanced neural network and backpropagation algorithms to boost detection and classification accuracy. Model generalizability will be enhanced through the adoption of transfer learning and data augmentation techniques (e.g., rotation, scaling), enabling performance adaptation to other arid regions and diverse archaeological contexts. Establishing a framework for long-term multimodal investigations will also help generate reliable, labeled training datasets to further refine and guide excavation strategies.

Data Collection and Field Validation:

Extended field validation will be essential to confirm and fine-tune the predictions generated through remote sensing and machine learning. Planned field verification and ground-truthing surveys will ensure alignment with archaeological evidence. In addition, supporting geophysical methods—such as GPR and magnetic surveys—will be integrated to enhance subsurface validation. Further data collection will also include acquiring updated satellite imagery to support long-term site monitoring and improve the quality of training datasets.

Technology and Sensor Upgrades:

The project will benefit from expanded access to advanced remote sensing instruments, including thermal sensors and GPR, to complement the existing SAR and multispectral datasets. Integration of multitemporal and multisensor data sources will improve the detection of seasonal and environmental variations, enhancing the overall robustness of the prediction models.

Scalability and Regional Expansion:

The methodology developed at the Saruq Al-Hadid site will be adapted for application in similar arid environments across the Arabian Peninsula and globally. This scalability will help increase the impact of the solution and expand its contribution to the understanding and preservation of archaeological landscapes in desert and environmentally challenging settings.

Interdisciplinary and Institutional Collaboration:

Future work will continue to foster collaboration between remote sensing experts, archaeologists, and machine learning specialists to align technological development with heritage conservation priorities. Expanding partnerships with academic institutions and research centers in related fields will strengthen the scientific basis of the work. Collaborations with technology providers and data agencies can support access to up-to-date, high-quality imagery and remote sensing resources.

## 3 Use Case Requirements

**REQ-01**: High-Resolution Imagery Acquisition
Must obtain high-resolution multispectral imagery (e.g., Worldview-3) to capture detailed surface features.

**REQ-02**: SAR Data Integration
Must incorporate SAR data (e.g., ALOS-2/PALSAR-2) to enable subsurface feature detection in arid environments.

**REQ-03**: Advanced Image Processing Capabilities
Must support advanced techniques such as PCA and spectral band transformations to enhance and extract relevant features.

**REQ-04**: Geospatial Data Integration and Analysis
Must integrate geospatial datasets (e.g., digital elevation models, slope maps, hydrographic networks) and perform spatial analysis using GIS tools.

**REQ-05**: Machine Learning and Deep Learning Infrastructure
Must include infrastructure for unsupervised classification (e.g., ISODATA, K-means++) and

deep learning (e.g., U-Net) for semantic segmentation, supported by high-performance computing (e.g., GPU clusters).

**REQ-06**: Multisensor and Multitemporal Data Handling
Must process and fuse data from multiple sensors and across different time periods to enhance robustness and accuracy.

**REQ-07**: Field Validation and Ground-Truth Integration
Must include mechanisms to collect and integrate ground-truth data (e.g., GPS surveys, excavation records) to validate and refine predictions.

**REQ-08**: Data Storage, Processing, and Management
Must support scalable data storage and management of large volumes of remote sensing and derived datasets.

**REQ-09**: Use of Established Software Tools
Should use established platforms such as ArcGIS Pro and Environment for Visualizing Images (ENVI)/Interactive Data Language (IDL)for processing, modeling, and visualization.

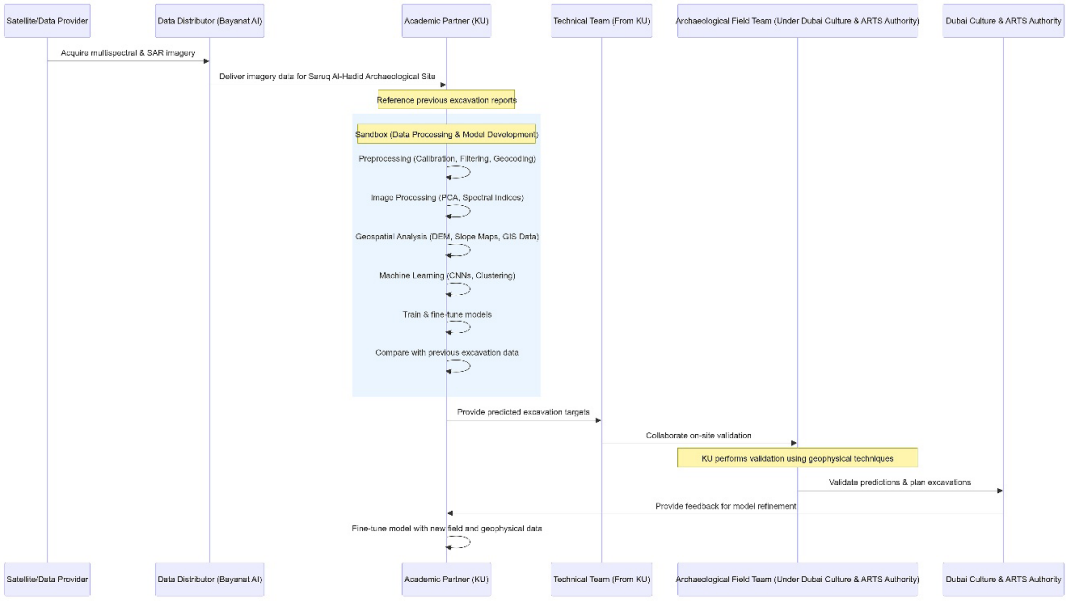
**REQ-10**: Scalability and Adaptability
Should be scalable to larger study areas and adaptable to other desert or arid environments.

**REQ-11**: Multidisciplinary Expertise
Must involve a team with expertise in archaeology, geospatial analysis, and artificial intelligence for system development and interpretation.

**REQ-12**: User Interface and Visualization Tools
Should include user-friendly visualization tools to support archaeologists and decision-makers in interpreting results and guiding fieldwork.

**REQ-13**: Geophysical Validation Integration
Must support the incorporation of ground-truth geophysical data (e.g., GPR, magnetic survey) to validate and enhance remote sensing-based predictions.

## 4 Sequence Diagram




## 5 References

[1] Ben-Romdhane, H.; Francis, D.; Cherif, C.; Pavlopoulos, K.; Ghedira, H.; Griffiths, S. (2023). Detecting and Predicting Archaeological Sites Using Remote Sensing and Machine Learning–Application to the Saruq Al-Hadid Site, Dubai, UAE. Geosciences, 13, 179. https://doi.org/10.3390/geosciences13060179

[2] Santos-Assunção, S., Ali, M., Francis, D., Ali, M.Y., Cherif, C., Radwan, M.B., Alsuwaidi, M.A., Al Ali, B.M., Safi, J.H., Iqbal, M.A. and Ullah, S., Uncovering Archaeological Treasures at Saruq al-Hadid, UAE: Insights from Ground Penetrating Radar and Magnetic Data, (2025) in press, Science of Remote Sensing Journal, Elsevier.

[3] Ali, M., Santos-Assunção, S., Ali, M.Y., Francis, D., Radwan, M.B., Alsuwaidi, M.A., Al Ali, B.M., Ullah, S. and Iqbal, M.A. (2025). Uncovering Archaeological Treasures at Saruq al-Hadid, UAE: Insights from Ground Penetrating Radar and Magnetic Data. under review, Khalifa University of Science and Technology & Dubai Culture and Arts Authority.

# Use Case 4: Artificial Intelligence + Robots Reconstruct Parking Space and Solve Urban Parking Problems

**Organization:** Shanghai Artificial Intelligence Research Institute

**Country:** China

**Contact Person:**

WANG Yijing (Ms), wangyijing@sairi.com.cn

## 1 Use Case Summary Table

| Item | Details |
|---|---|
| Category | Smart home/cities |
| Problem Addressed | Clearly describe the primary issue or challenge |
| Key Aspects of Solution | List main elements of the proposed solution, including the role of AI. |
| Technology Keywords | Relevant technical keywords |
| Data Availability | Public / Private (specify sources or repositories if available) |
| Metadata (Type of Data) | Specify data types (e.g., text, audio, visual) |
| Model Training and Fine-Tuning | Briefly describe AI/ML technologies and approaches used |
| Testbeds or Pilot Deployments | Provide links or references to deployment examples |
| Code repositories | Provide github links (if available) |

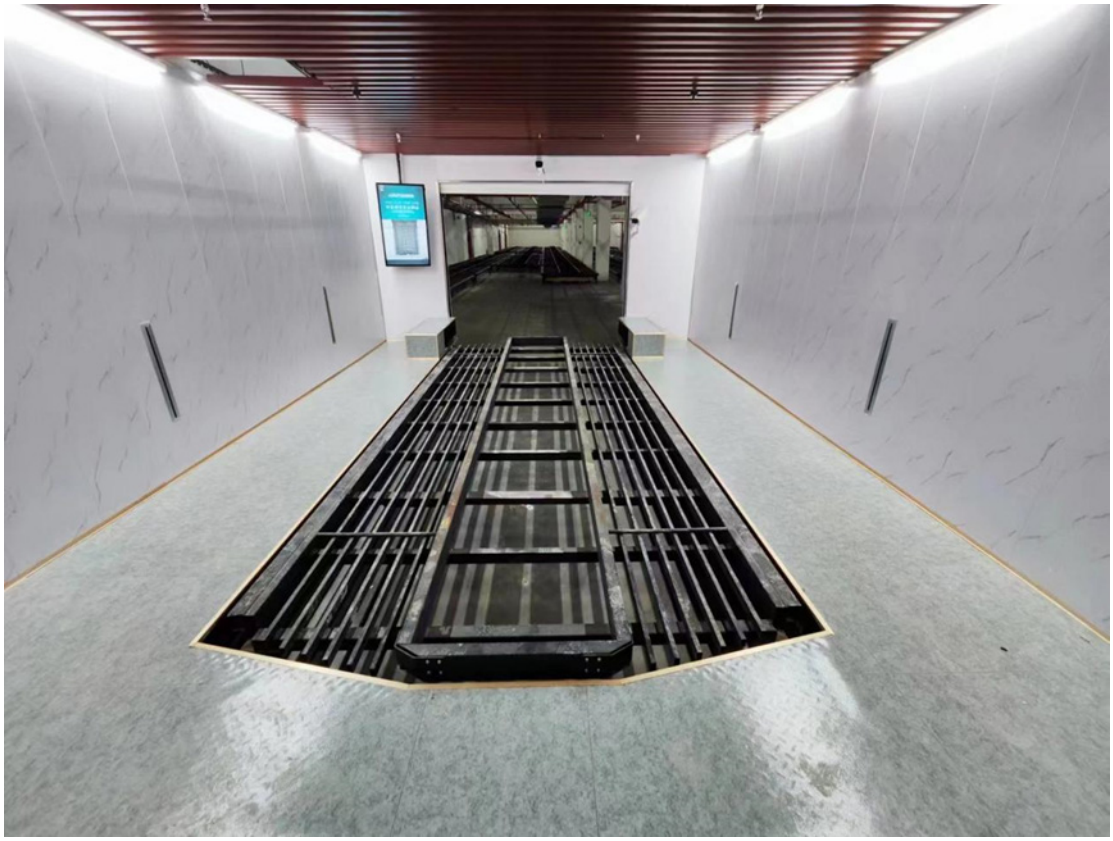
## 2 Use Case Description

### 2.1 Description

At present, artificial intelligence robots have been applied in parking garages. They closely combine artificial intelligence and big data technologies such as intelligent perception, planning algorithms, and multi-concurrent scheduling, and apply the fully automatic parking robot system with intelligent perception and motion control to the existing parking garages, solving the problems of parking space shortage, difficulty in finding parking spaces, and difficulty in finding cars in existing parking lots. Through full-process intelligent operation, the previous parking experience is completely overturned. By building and renovating intelligent parking lots, the parking improvement rate can reach 60%-80% when a single-layer parking solution is adopted; when a double-layer parking solution is adopted, the parking improvement rate

can reach more than 110%, saving single-parking space, increasing parking spaces, improving parking efficiency, and saving labor costs. There are also multiple projects under construction.

Before the AGV Robot assumes control of the vehicle, a 3D scanning system at the entrance conducts a comprehensive dimensional inspection, including height measurement. If the vehicle exceeds the maximum allowable height limit, the AGV Robot will abort the transportation process. Simultaneously, the system triggers an automated warning notification to alert operators of the oversize condition. Two radar sensors are positioned at opposite corners of the AGV Robot, enabling 360-degree environmental monitoring. If a human or obstacle is detected within its path, the AGV Robot will automatically halt or adjust its trajectory to ensure safe navigation. Additionally, the robot's operational zone is strictly segregated from human activity as part of critical safety protocols.



## 2.2 Benefits of use case

With the continuous improvement of urbanization, the per capita car ownership continues to grow rapidly, and the contradiction between supply and demand of urban parking lots is becoming increasingly acute. A large number of parking lots built in the early stage cannot meet the current parking needs in terms of planning and design.

Relevant data show that there is a shortage of 80 million parking spaces in China. The shortage of parking spaces in large cities around the world is also a common urban governance problem. The difficulty of parking not only affects the daily quality of residents, but also brings many obstacles to the city's commercial activities, public services, etc., resulting in a huge waste of time costs and a significant reduction in urban operation efficiency.

The existing parking lots lack effective technologies and products to increase the number of parking spaces, which is the main contradiction faced in solving the parking problem in the current urban renewal. Intelligent technologies such as self-service payment systems and parking guidance systems can better improve the parking and car-picking experience, but cannot effectively increase the number of parking spaces, and mechanical stereo garage technology is mainly used in new projects.

In this context, the combination of robots and artificial intelligence technologies is seen as an effective way to solve this problem. Without the need for construction and renovation and multi-story construction, the renovation of an ordinary parking lot with about 100 parking spaces can be completed within 1 month. After the upgrade, the number of parking spaces can be increased by about 50%, and the annual return on investment can reach 30%.

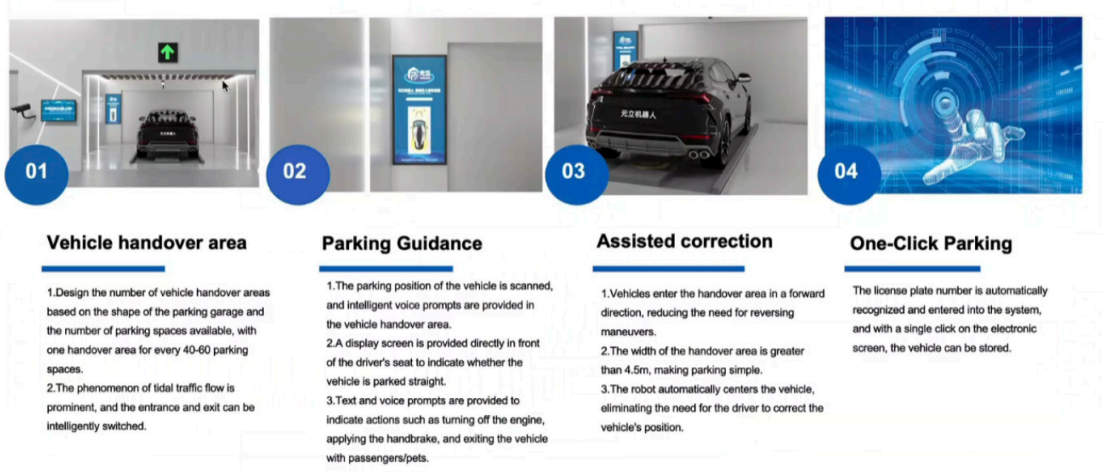
**Partners**

Shanghai Jiao Tong University Robotics Institute, Parallel Computing Institute, Beidou Navigation Institute, School of Cyberspace Security, Shenzhen Artificial Intelligence Industry Association, China Mobile Robot Industry Alliance

## 3 Use Case Requirements

- **REQ1**: it is critical that vehicles handover areas are designed based on the shape of the parking garage.
- **REQ2**: it is critical that parking guidance is provided to guide the driver in the vehicle handover area
- **REQ3**: it is expected that robotic assistance is provided to correctly position the vehicle in the handover area.
- **REQ4**: It is critical that the vehicle parking is indexed by vehicle license plate number.

## 4 Sequence Diagram



## 5 Future Work

## 6 References

[1] **Car2V.** *Car2V Vehicle Communication Platform.* Accessed June 19, 2025. https://www.car2v.com.

[2] **Shanghai Artificial Intelligence Industry Association (SAIIA).** *About Us.* Accessed June 19, 2025. https://www.saiia.org.cn/index.php/1-2/about/.

[3] **Shanghai Jiao Tong University, EPCC.** *Electronic Packaging and Circuit Center.* Accessed June 19, 2025. http://epcc.sjtu.edu.cn/.

[4] **AGV-AMR.** *Industry Leadership.* Accessed June 19, 2025. http://www.agv-amr.com/ldlm/list.php?catid=25.

[5] **Belt and Road Development Institute.** *Research Projects.* Accessed June 19, 2025. http://www.bdi.org.cn/page3.

[6] **Shanghai Jiao Tong University, Information Security Center.** *About Us.* Accessed June 19, 2025. https://infosec.sjtu.edu.cn/About.aspx.

# Use Case 5: AI-Powered Video Ringback Tone

**Organization**: iMUSIC Culture & Technology Co.,Ltd.

**Country**: China

**Contact Person(s)**:

An Lin, lina.ua@chinatelecom.cn

Guanzheng Xiao, xiaogz.ua@chinatelecom.cn

## 1 Use Case Summary Table

| Item | Details |
|---|---|
| Category | Smart home/cities |
| Problem Addressed | Pre-call information and transmission issues, users' self-expression needs, merchants' product promotion needs |
| Key Aspects of Solution | Model training and fine-tuning constitute core processes in AI system development. Supervised learning typically drives model training, involving data preprocessing (cleaning, normalization, augmentation), architecture selection (e.g., Transformer), loss functions (cross-entropy), and optimizers. Regularization techniques (Dropout, L2) combined with early stopping and cross-validation prevent overfitting. Distributed training and mixed-precision computing accelerate large-scale model development.<br>Fine-tuning leverages pretrained models (e.g., Bidirectional Encoder Representations from Transformers(BERT), Residual Network(ResNet)) through transfer learning: freezing lower-layer parameters while adapting top layers for new tasks. It employs reduced learning rates (typically 1/10 of initial training) with category-specific data. Advanced techniques include layer-wise learning rate adjustment (lower rates for base layers), knowledge distillation, and adapter module insertion. Data augmentation and domain adaptation methods (e.g., adversarial training) enhance cross-domain generalization. Performance validation utilizes hold-out test sets and metrics (accuracy, F1-score). |
| Technology Keywords | Diffusion, transformer, Generative Adversarial Networks(GAN), Retrieval-Augmented Generation(RAG), Diffusion Transformer(DiT) |
| Data Availability | Private |
| Metadata (Type of Data) | Text, audio, image |

(continued)

| Item | Details |
|---|---|
| Model Training and Fine-Tuning | The AI video ringtone product leverages AI-Generated Content (AIGC) technology to create personalized music short videos of approximately 20 seconds in length based on user input, which can include text or images. This technology involves several key AI/ Machine Learning approaches to ensure the generated content is both engaging and relevant. |
| Testbeds or Pilot Deployments | There are currently no testbeds or pilot deployments mentioned in the use case. The project is in a preparatory phase, with no stated collaborations or trial environments initiated at this time. |

# 2 Use Case Description

## 2.1 Description

Before the phone is connected, a certain connection cannot be established between the caller and the called person.AI video ringback tones are used to solve the above problems.

Users hope to convey certain information to others in advance before making a phone call, such as their professional identity, beautiful photo image, talents, daily life, favorite music videos, corporate advertisements, etc., which need to be presented in the form of videos.Therefore, we have launched the AI video ringtone product.

Based on AIGC technology, generate a music short video of about 20 seconds from user input text or images.

AI video ringback tones make information transmission in mobile communication more efficient and valuable. At the social level, promoting the rapid dissemination of information, breaking geographical limitations, allowing remote areas to access cultural and entertainment information, and also transmitting their own information to the outside world. From an environmental perspective, it reduces unnecessary travel caused by information exchange, lowers transportation energy consumption and exhaust emissions.

**Use Case Status:**

When the calling user initiates a call, the system feature enters the administrative state of Ready, indicating that the basic call capability is ready. However, because the video ringtone playback has not yet been triggered, the execution status is marked as Not Run. The critical link occurs when the system triggers and attempts to play the video ringtone.

The critical link occurs when the system triggers and tries to play the video ringtone (the called network requests the video from the platform and pushes it to the called terminal). The administrative state can be Ready, but if the ringtone is not available due to, for example, cross-operator network incompatibility, it will go to Block, meaning that the function is temporarily unavailable. Correspondingly, the execution status reflects the actual result of the step: if the ringtone is successfully acquired and ready to be played, it is marked as Passed; if the content is invalid, the platform interface is wrong or loading fails, it is marked as Passed. If the content fails, the platform interface is wrong, or the loading fails, it is labeled as Error or Failed.

At the stage when the video content is actually played at the called terminal, the core function should remain in the administrative state of Ready, while the execution state provides direct feedback on the playback result: A smooth display of a 20-second video is Passed; if playback fails due to terminal incompatibility, lag or interruption, it is Failed. Failed.

The next part of the called user's response (answer, reject or no action) does not involve the management state, but purely user behavior triggers different execution states The next part of the called user's operation response (answer, reject or no operation) does not involve the management state, it is purely the user's behavior that triggers the different execution states: the user takes the initiative to answer, the system successfully establishes the connection, which marks the achievement of the main business objectives, and the execution state is Passed; the user's refusal to answer results in the caller receiving If the called user does not operate within the ringing timeout period, the system automatically ends the call, and the execution state can be regarded as Skip because this interaction is not handled effectively.

Finally, in the final part of call establishment or termination, the successful establishment of the call belongs to Passed state, while the failure of the call due to rejection or timeout is a Failed state. The administrative state is not usually applied directly here.

**Partners:** N/A

## 2.2 Benefits of the use case

AI video ringback tones make information transmission in mobile communication more efficient and valuable. At the social level, promoting the rapid dissemination of information, breaking geographical limitations, allowing remote areas to access cultural and entertainment information, and also transmitting their own information to the outside world. From an environmental perspective, it reduces unnecessary travel caused by information exchange, lowers transportation energy consumption and exhaust emissions.

## 2.3 Future Work

On the one hand, we focus on user needs and provide services such as commercial promotion, personal display, and life tips to users who have the ability to transmit information before a call;

On the other hand, we will continue to explore the value of AI technology, promote its implementation, create products, and benefit more users.

# 3 Use Case Requirements

**REQ-01**: Multi-Modal Input Support

This requirement calls for The system has the ability to handle heterogeneous input data. Users can use text descriptions (≤100 characters of natural language) or image files (Joint Photographic Experts Group/Portable Network Graphics(JPG/PNG) format, resolution ≥1080P, size ≤5MB) as the video generation source. The key technology implementation needs to deploy multimodal alignment model (e.g., Contrastive Language–Image Pretraining(CLIP) architecture) to map textual semantics and visual features to a unified implicit space, to ensure that the input content parsing delay is ≤3 seconds (P95 response time), and the parsing results

need to be outputted in structured data format (JSON schema with keywords/subject labels and confidence level).

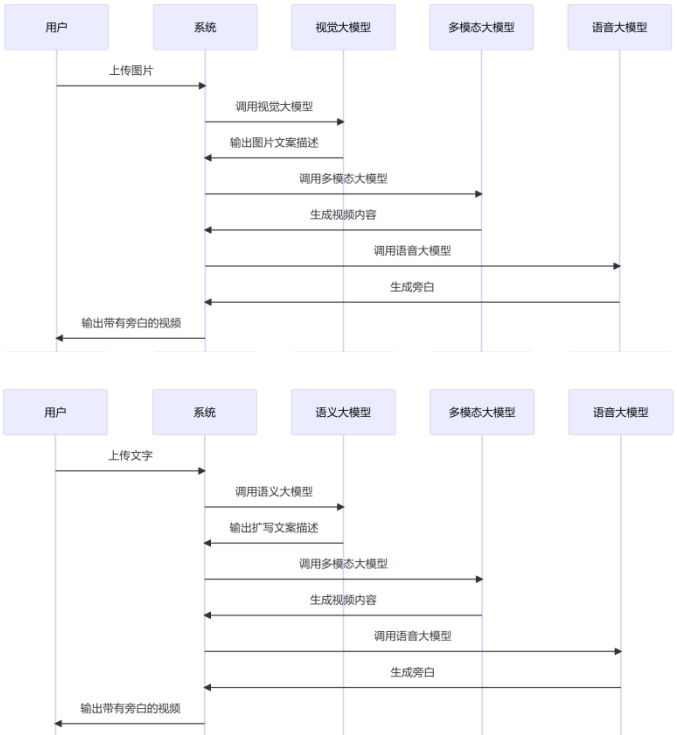
**REQ-02**: High Fidelity Motion Video Generation[1]

The system is required to generate 20 seconds of video (720P@25fps, H.264 encoding) based on the Diffusion+Diffusion Transformers(DiT) fusion architecture to realize an end-to-end generation pipeline[2]. Technical requirements include:Layered diffusion strategy (Latent Diffusion) to reduce computational complexity.Optical Flow Long Short-Term Memory(LSTM) to ensure inter-frame coherence (Structural Similarity Index Measure(SSIM) ≥ 0.85) .Integration of audio-video synchronization controllers (based on Mel-Frequency Cepstral Coefficients(MFCC) feature alignment) .Generation process to meet the constraints of real-time (end-to-end delay of ≤ 10 seconds / piece), the output content should be Pass automated validation (content relevance score ≥ 90%, evaluated by Tencent Video Multimethod Assessment Fusion(VMAF) tool).[3]

**REQ-03**: Operator-level Deployment Compatibility

The solution must be compatible with the technical specifications of the video ringtone platforms of China's three major operators (China Telecom/Mobile/Unicom), and the core capabilities include: Dynamic transcoding engine: support for the conversion of the input video to the target format (e.g., China Mobile's H.265 2Mbps bit rate). Terminal adaptation layer: realize Android 9+ (ExoPlayer Software Development Kit(SDK)) and iOS 13+ (Audio-Video(AV) Foundation) native playback support. Quality of Service(QoS) guarantee mechanism: first frame loading time ≤ 1.5 seconds (under Round-Trip Time(RTT) 100ms network environment), frame loss rate ≤ 0.1%.

## 4 Sequence Diagram



## 5 References

[1] High-Resolution Image Synthesis with Latent Diffusion Models, https://arxiv.org/abs/2112.10752

[2] Attention Is All You Need, https://arxiv.org/abs/1706.03762

[3] Scalable Diffusion Models with Transformers, https://arxiv.org/abs/2212.09748

# Use Case 6: YI JIANG JIE: Cultural and Creative Explanation of Big-Scale Model

**Organization**: China Telecom Wanwei Information Technology Co., Ltd.

**Country**: China

**Contact Person(s)**:

Junru.Chen, 644211755@qq.com

Yuanxin.He, 1540016182@qq.com

Zuoji.Chen, 2469885370@qq.com

## 1 Use Case Summary Table

| Item | Details |
|---|---|
| Category | Smart home/cities |
| Problem Addressed | Inefficient Service Delivery, Monotonous Experience Design, Limited Cultural Engagement, High Operational Costs |
| Key Aspects of Solution | The proposed solution involves using AI to provide intelligent private tour guide services for tourists in museums and scenic spots, enhance their visiting experience by facilitating easy acquisition of historical and cultural knowledge, explore and showcase the connotations of heritage in a digital and intelligent way, enable rapid content updates to keep tourists informed about the latest cultural trends, encourage questions and exchanges through multimodal interactions, and break the limitations of time and space for efficient cultural dissemination. |
| Technology Keywords | Large Language Model, Retrieval-Augmented Generation (RAG), Multimodal Agent |
| Data Availability | Private |
| Metadata (Type of Data) | Text, audio, image |
| Model Training and Fine-Tuning | The model training and fine-tuning of the Yi Jiang Jie platform is centered around a domain-specific, multimodal AI system tailored to the needs of cultural heritage interpretation. It leverages Large Language Models (LLMs), Retrieval-Augmented Generation (RAG), and multimodal agents, with focused training on museum- and heritage-specific content. |
| Testbeds or Pilot Deployments | The testbeds or pilot deployments of the Yi Jiang Jie system likely involve early-stage, real-environment implementations that validate technical feasibility, usability, and cultural relevance before full-scale rollout. China Telecom Museum, Gansu Provincial Museum |

# 2 Use Case Description

## 2.1 Description

As the global culture and tourism industry accelerates its digital transformation, traditional cultural heritage interpretation services face four critical challenges:

Inefficient Service Delivery: Manual guiding is constrained by manpower limitations, leading to supply-demand imbalances during peak seasons and inconsistent service quality due to varying expertise levels among guides. This model struggles to scale, leaving visitors in high-traffic venues underserved and dissatisfied.

Monotonous Experience Design: Traditional hardware-based audio guides suffer from fixed, rarely updated content and a lack of interactivity, failing to engage modern audiences. Their cumbersome rental/return processes and limited functionality (e.g., no real-time updates or personalized responses) create friction in the user journey, resulting in subpar experiences.

Limited Cultural Engagement: Current services rely on one-way information delivery, neglecting the younger generation's demand for immersive, personalized interactions. This static approach fails to contextualize heritage stories or adapt to visitor interests, leading to low knowledge retention and missed opportunities for meaningful cultural connection—with less than 50% of exhibit information effectively absorbed by visitors.

High Operational Costs: Maintenance of traditional equipment incurs significant expenses due to frequent wear-and-tear, manual inventory management, and technical support needs. These costs escalate further in large institutions, straining budgets and diverting resources from core heritage preservation efforts.

In order to enhance the visiting experience of tourists in museums and scenic spots, address the shortcomings of traditional guided tour services, make full use of artificial intelligence technology, and promote the innovative development of intelligent services for cultural heritage, China Telecom Gansu Company has planned the AI Yi Jiang Jie product based on artificial intelligence.

The Yi Jiang Jie product aims to provide intelligent private tour guide services for tourists in museums and scenic spots, helping them easily acquire historical and cultural knowledge during the visit and enhancing their visiting experience. At the same time, as an innovative application of intelligent services for cultural heritage, it explores and showcases the connotations of heritage in a digital and intelligent way. It can not only update the content rapidly, enabling tourists to keep abreast of the latest cultural trends and expand their knowledge, but also encourage questions and exchanges through multimodal interactions. The convenient access also breaks the limitations of time and space, enabling the efficient dissemination of culture both inside and outside the museums and scenic spots.

Yi Jiang Jie is constructed and implemented based on the architecture of "Large Language Model(LLM) + Retrieval-Augmented Generation (RAG) + Multimodal Agent". Driven by technological innovation, the platform creates a multimodal intelligent service application.

Relying on the deeply integrated architecture of "Semantic Vector Model + Large Model + Knowledge Base + RAG", we've innovatively developed an intelligent Q&A system. The cross-venue semantic vector model enables precise analysis of visitors' question intents and

generation of contextualized responses. Combined with the DeepSeek-R1 large model, it generates professional interpretive content. Leveraging the pre-built dynamic knowledge base and RAG, we ensure a dual improvement in the authority and naturalness of answers. It also supports continuous multi-round in-depth interactive Q&A and personalized expression adjustments. Moreover, by integrating visitors' behavior data to build profiles, it can achieve intelligent adaptation for exhibit recommendations and the depth of explanations.

We've innovatively created a "voice + text + photo + location" multimodal interaction system, integrating multimodal capabilities such as Automatic Speech Recognition (ASR), Text-to-Speech (TTS), visual recognition, and semantic understanding. Through the RAG framework, we achieve the coordinated operation of knowledge retrieval and response generation.

**Use Case Status:** The use case is in progress and part of a larger product development.

### Partners

China Telecom Museum, Gansu Provincial Museum, Zunyi Conference

## 2.2 Benefits of the use case

1. Elevating Visitor Experience: Intelligent, Immersive Exploration

Yi Jiang Jie delivers an interactive AI-guided experience that mimics a personal tour guide, adapting to real-time visitor needs:

Contextual, Real-Time Engagement: Leveraging real-time positioning technology, the platform automatically triggers exhibit explanations as visitors approach, enabling seamless, on-demand knowledge access.

Innovative Cultural Interaction: Features like AI Cultural Pass (gamified heritage exploration) and Heritage Cosplay Experience (virtual role-playing with historical figures) blend education with entertainment, deepening emotional connections to cultural stories.

Tailored to Curiosity: Supports natural-language Q&A, multi-turn dialogues, and personalized route recommendations, ensuring every visitor—from casual sightseers to history enthusiasts—finds relevance in their journey.

2. Empowering Cultural Institutions: Digital Transformation Partner

For museums and heritage sites, Yi Jiang Jie serves as a strategic technology ally:

Knowledge Asset Management: Builds institution-specific large models to centralize cultural relic data, academic research, and visitor insights, enabling sustainable knowledge asset management and cross-departmental collaboration.

Data-Driven Optimization: Analyzes user interaction data (e.g., popular exhibits, query trends) to provide actionable insights for exhibition layout redesign, content prioritization, and visitor flow optimization.

Research & Academic Support: In terms of scientific research, it assists the museum in applying for project topics and creating academic papers, thereby enhancing the museum's academic influence and generating more economic and social benefits.

3. Leading Industry Innovation: Redefining Heritage Engagement

Yi Jiang Jie pioneers next-generation heritage experiences to future-proof the industry:

Multimodal Storytelling: Yi Jiang Jie transforms the explanations of cultural relics into interesting historical stories, interactive 3D models, etc., and enhances the effectiveness of cultural popularization in novel forms.

Phygital (Physical+Digital) Synergy: Develops augmented reality (AR) guided tours and virtual exhibitions, merging physical artifacts with digital enhancements to create scalable, low-cost engagement tools for institutions of all sizes.

Standard-Setting Impact: By demonstrating how AI can streamline operations, personalize experiences, and drive cross-sector collaboration, Yi Jiang Jie sets a global benchmark for intelligent heritage services, inspiring peers to adopt data-driven, visitor-centric models.

4. Fostering Cultural Prosperity: Beyond Tourism to Legacy

The platform's impact extends to broader cultural and economic ecosystems:

Cultural Commerce Activation: Yi Jiang Jie attracts more people to pay attention to cultural tourism, drives the consumption of cultural and creative products in and around museums, and promotes the development of the cultural industry.

Global Reach, Local Roots: Breaks through time and space barriers via 24/7 mobile access, enabling pre-visit research, remote exhibitions, and multilingual support (under development). This makes China's cultural heritage a shareable, universal experience—fostering national identity and cross-cultural understanding.

## 2.3 Future work

1. Specialized Cultural Heritage Intelligence

Leveraging the Cultural and Creative Explanation of Big-scale Model, we focus on "intelligent cultural heritage services" to elevate AI-driven interpretation. This enables scenario-specific explanations for indoor exhibitions, outdoor grottoes, and scenic spots, aligning with the unique characteristics of each environment—delivering tailored experiences that respect architectural, historical, or natural context.

2. Age-Inclusive Design

We craft age-adaptive experiences by analyzing behavioral patterns, cognitive levels, and interests of visitors across generations—senior, adult, young, and juvenile. Whether simplifying language for families, deepening historical narratives for scholars, or incorporating interactive elements for youth, our solutions ensure relevance and engagement for all.

3. Diverse Style Customization

Catering to varied preferences, we offer multi-style interpretation options, including:

- General: Accessible overviews for casual visitors

- Engaging: Story-driven, interactive content for broad appeal
- In-Depth: Academic insights for enthusiasts
- Extended: Supplementary context for deeper exploration

This flexibility allows users to choose experiences that match their curiosity and visit goals.

4. End-to-End R&D Pipeline

Our process includes:

- Data Collection: Curating multilingual, multimodal datasets to enrich model training
- Proof-of-Concept Development: Rapid prototyping to validate technical feasibility in real settings
- Model Refinement: Iterative improvements to enhance accuracy, fluency, and scenario adaptability
- Standardization: Developing industry-aligned protocols for data security, model interoperability, and service quality

These efforts not only drive product innovation (new features, expanded functionality) but also set benchmarks for AI in cultural heritage, positioning us as both practitioners and standard-bearers.

## 3 Use Case Requirements

**REQ-01**: Multimodal Interaction & Intelligent Language Capabilities

Yi Jiang Jie integrates natural language processing (NLP), image recognition, and speech recognition technologies to deliver rich visual and auditory experiences for visitors. Its cultural heritage large model—trained via deep learning on massive historical and cultural datasets—enables intelligent agents to understand and generate human language, precisely parsing user queries and delivering context-aware responses. This showcases exceptional interactivity and adaptability, ensuring seamless engagement across diverse visitor needs.

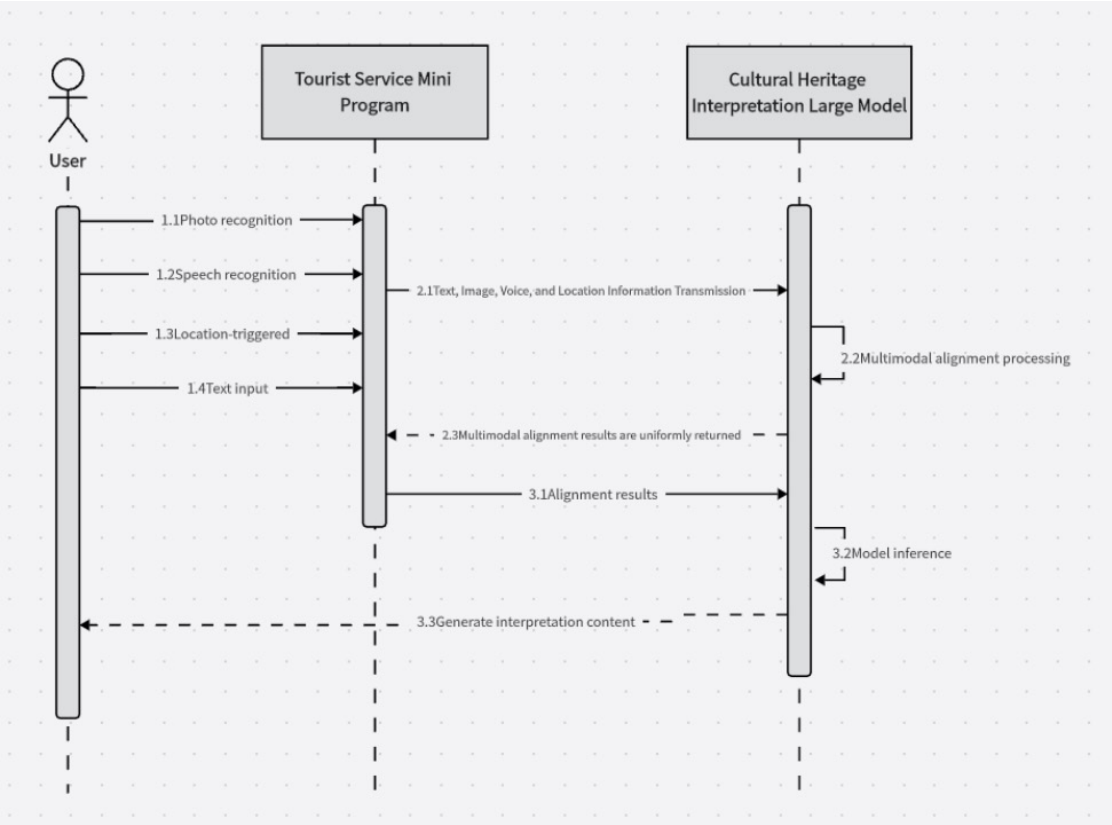
**REQ-02**: Accessible Mini-Program Form Factor

Built as a mini-program, the product eliminates the need for physical device rental or return, offering effortless access and updates. This design drastically lowers user access barriers while reducing museums' operational and maintenance costs—making advanced AI guidance scalable and cost-effective for cultural institutions of all sizes.

**REQ-03**: Real-Time Personalization & Continuous Improvement

Yi Jiang Jie's real-time positioning and personalized service engine tailors experiences to visitors' specific needs and behavioral patterns, delivering custom routes, exhibit recommendations, and explanation depth. The platform also features user feedback functionality, allowing visitors to rate content and suggest improvements—directly informing iterative model optimizations to refine explanatory accuracy and relevance.

## 4 Sequence Diagram

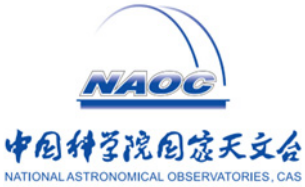



## 5 References

[1] China Telecom Museum. (n.d.). Collection catalogs, research reports, exhibition descriptions [Online]. Available at: http://museum.chinatelecom.com.cn

[2] Gansu Provincial Museum. (n.d.). Collection catalogs, research reports, exhibition descriptions, and academic papers [Online]. Available at: http://www.gansumuseum.com/home

# Use Case 7: OneAstronomy[1]: Harnessing Multimodal Astronomical Foundation Models to Pave a New Path for Humanity's Sustainable Development




**Organization**: Zhejiang Lab, National Astronomical Observatories of China (NAOC)

**Country**: China

**Contact Person(s)**:

Yangbin Xie xieyangbin@zhejianglab.org

Dandan Huang huangdandan@zhejianglab.org

Wanqi Sun sunwanqi@zhejianglab.org

## 1 Use Case Summary Table

| Item | Details |
|---|---|
| Category | Smart home/cities |
| Problem Addressed | OneAstronomy is expected to address the following challenges:<br>1. Technological barriers in exploring cosmic origins and identifying habitable exoplanets.<br>2. Prediction and mitigation of solar activity impacts.<br>3. Inequities in scientific research and educational resource distribution. |
| Key Aspects of Solution | 1. Artificial Intelligence (AI)-driven Scientific Interpretation: OneAstronomy leverages Large Language Models (LLMs) and Multimodal AI to interpret complex astronomical data [2][3][4], enabling breakthroughs in the understanding of cosmic origins and the identification of habitable exoplanets.<br>2. Solar Activity Monitoring: Through unified encoding of different modalities, OneAstronomy uses AI to forecast solar activity and assess potential impacts, contributing to early warning systems and space weather mitigation.<br>3. Multilingual and Inclusive Access: OneAstronomy supports multi-language interaction, making advanced astronomical tools accessible to global users regardless of language or region.[5]<br>4. Fairness and Bias Mitigation: The system incorporates bias mitigation techniques to reduce cultural, gender, and regional biases, promoting equity in scientific research and education.[6] |
| Technology Keywords | Large Language Models (LLM), Multimodal AI, Contrastive Language-Image Pretraining (CLIP), Knowledge Distillation, Bias Mitigation, Multi-language |

(continued)

| Item | Details |
|---|---|
| Data Availability | Private<br>The dataset is planned to be gradually opened in the future, with the specific release schedule yet to be determined. |
| Metadata (Type of Data) | Text Data, Image Data, Multimodal Data Image-Text Pairs , Time Series Data, Spectral Data |
| Model Training and Fine-Tuning | OneAstronomy is trained on large-scale, high-quality astronomical datasets. On top of a strong language foundation model, we apply a multi-stage fine-tuning process including Continual Pretraining (CPT), Supervised Fine-Tuning (SFT), and Direct Preference Optimization (DPO), to enhance domain-specific capabilities in astronomy.[3][[7]<br>In parallel, we build dedicated foundational models for solar imagery and time-domain observations using self-supervised learning. These models are then aligned with the language model using CLIP-style contrastive learning to form a unified multimodal system.[8][9][10]<br>To ensure practical deployment and inclusive application, the training pipeline also includes techniques such as knowledge distillation, bias mitigation for fairness, and multilingual enhancement to support broader accessibility.[5] |
| Testbeds or Pilot Deployments | https://OneAstronomy.zero2x.org/ |
| Code repositories | OneAstronomy has been made available for free use globally. The model is also planned to be open-sourced on GitHub in the future, as part of the Zhejiang Lab's Open Science Initiative. |

# 2 Use Case Description

## 2.1 Description

Humanity's pursuit of high-quality sustainable development not only requires further exploration of space to discover exploitable extraterrestrial resources but also necessitates precise predictions of various celestial activities, such as the Sun, to avoid potential disaster impacts. However, astronomical research faces challenges such as the diverse and complex modalities of data and high barriers to entry for study. Current astronomical research is undergoing a paradigm shift from hypothesis-driven to data-driven approaches. Artificial intelligence has become a key enabler for exploring the universe.[4][11] Therefore, we developed the OneAstronomy series of models based on Qwen2.5.[12]

OneAstronomy maintains the question-answering dataset in the following ways:

1) Pre-training question-answering data for general large models: Used for foundational language capability learning.
2) Retrieval-augmented Generation (RAG) knowledge base data: Provides specific knowledge support in the field of astronomy.
3) Real-time internet search data: Captures dynamic and time-sensitive information.

4) High-quality annotated data based on reasoning and fine-tuning: Used for model fine-tuning and reasoning optimization.

The training of the OneAstronomy model employs a two-stage methodology.

The first stage is Continual Pre-training (CPT), aimed at integrating diversified astronomical corpora

to establish a specialized knowledge foundation. Data for this stage is sourced from the public

RedPajama[13] astronomical subset (2.61 million records, ~4.31GB) and information from

Dolma[14] astronomical subset which is made available under the Open Data Commons Attribution

License (1.73 million records, ~2.2GB), as well as datasets internally curated with open accessible

astronomical papers with proper license (32,000 records, ~4.9GB).

The second stage is Supervised Fine-tuning (SFT), designed to enhance domain-specificinstruction

comprehension and execution capabilities by integrating expert-annotated astronomical question

answer pairs and inference-focused data. SFT data originates from public datasets including

curated subsets of open-community resources like ShareGPT_Vicuna_unfiltered[15] and

MetaMathQA[16] (covering diverse instructional tasks, yielding 3.50 million general instruction

records) and Multiple HuggingFace open-source platform datasets with permissive licenses

(yielding 360,000 general reasoning records), supplemented by internally curated datasets derived

from arXiv papers with proper licenses prior to March 12, 2025 (yielding 24,000 astronomical

reasoning records).[5]

Given the emphasis on fairness and bias mitigation, we have taken several targeted measures to ensure balanced dataset coverage. The current version primarily includes English and Chinese data, with plans to incorporate German, Spanish, French, and Portuguese in the next iteration. Specifically, ahead of the IAUS(International Astronomical Union Symposium) 397 conference in Greece, we evaluated OneAstronomy's performance on Greek through a dedicated set of assessment tasks, achieving results comparable to English. All data is sourced from publicly available global astronomy platforms, and analysis confirms comprehensive geographic representation, including contributions from researchers across Asia, North America, Africa, South America, and Oceania. While the dataset is focused on astronomical sciences, we have also evaluated the model's general capabilities, confirming that its performance remains consistent with the base model (Qwen2.5-72B [17]) and shows no evidence of bias toward or against other scientific fields.

OneAstronomy has the following capabilities:

(1) OneAstronomy was evaluated on two benchmark datasets. From the following charts, we can see OneAstronomy significantly outperforms its base model Qwen 2.5 in both astronomy knowledge and reasoning. It even surpasses GPT-4o [18] in astronomy-specific tasks. OneAstronomy significantly outperforms the peer astronomy-domain model AstroLLaMA 70B [[19][20] on the GPQA-astro [21] (Graduate-Level Google-Proof Q&A Benchmark, – Astronomy Subset) and NAOC-TianYi test sets.[4][5][22]

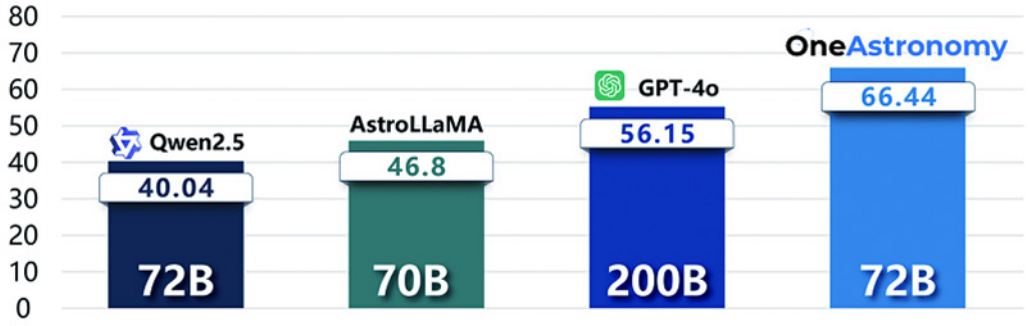
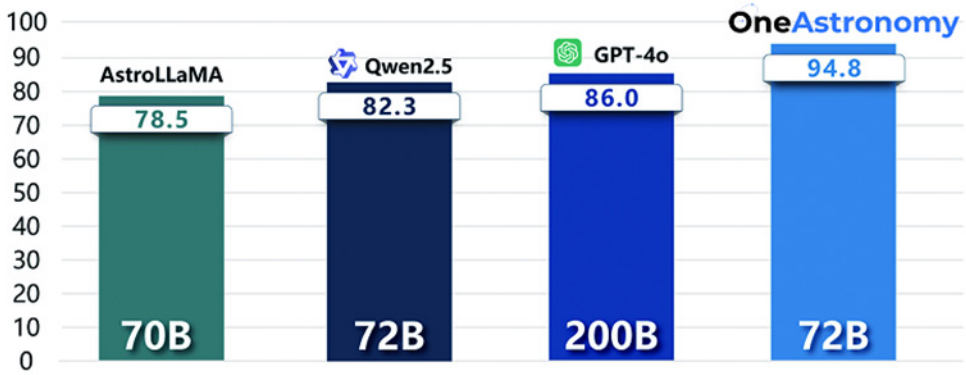



*Figure 1. OneAstronomy achieves a high score on the GPQA-astro test, an astronomical reasoning evaluation dataset with 447 questions. The score is calculated as the proportion of correctly answered questions (accuracy). The same metric is used in subsequent figures.*




(2) OneAstronomy possesses leading astronomical data analysis capabilities. For instance, stellar spectrum model can efficiently predict dozens of stellar parameters. It has identified over 8,000 candidates for extremely metal-poor stars, known as early universe messengers, whereas the total number identified by humans so far is merely around 50.[8]

(3) In addition, it also offers comprehensive capabilities for multimodal data integration, modeling, and prediction. For instance, the solar model currently under development leverages large-scale multimodal observational datasets—including documentation, spectral measurements, imagery, and time-series recordings—to build a multimodal, dynamically evolving foundation model. This model enables diverse downstream applications such as full-disk magnetic field forecasting at hourly and daily scales, Carrington-rotation–level predictions, and imputation of missing observations. By supporting end-to-end analysis of solar activity from historical data to future forecasts, it establishes a robust basis for solar hazard early warning, space-environment monitoring, and future deep-space exploration.[11]

**Use Case Status**

- The use case is part of a larger research project.
- The use case is part of a larger product development.

**Partners**

National Astronomical Observatories of China (NAOC)

OneAstronomy is jointly developed by Zhejiang Lab & NAOC . The main tasks of NAOC are to guide scientific research, provide domain knowledge, interpret and preprocess data, validate and improve the models.

## 2.2 Benefits of use case

The open intelligent platform and technological tools will break geographical limitations, allowing ordinary people in remote areas to access cutting-edge astronomical knowledge, research findings, and even telescope observation opportunities through a smartphone. This provides high-quality and equitable scientific research and educational resources, enabling everyone to explore the universe equally. Such initiatives contribute to building a shared future for mankind, fostering a global community that thrives on collective knowledge and mutual respect.

AstroOne can better predict and mitigate the impact of solar storms and other extreme space weather events on Earth's communication and navigation systems. Additionally, AstroOne can enhance the speed of near-Earth space warnings and improve monitoring capabilities for more distant spaces. This means it can provide faster alerts for potential threats in near-Earth space while also expanding our ability to observe and analyze conditions in deeper space, contributing to a more comprehensive space situational awareness.

AstroOne, through a census of over 200 million stars in the Milky Way, seeks out sun-like stars and Earth-like planets within our galaxy to aid humanity in the search for habitable worlds. By pushing the boundaries of existing observational technologies, AstroOne has mapped a spatiotemporal diagram of the Milky Way's evolution. This not only helps in understanding the galaxy's development but also delves into the early physical processes of the universe, pursuing the ultimate mystery of cosmic origins. This endeavor significantly advances our knowledge of space, offering insights that could lead to groundbreaking discoveries about the conditions necessary for life and the broader mechanisms governing the universe's structure and evolution.

## 2.3 Future Work

1. Optimization and Capability Enhancement of the Astronomical Foundation Model: Short-term priority will focus on improving the model’s reasoning and predictive accuracy by enriching the datasets and refining the modeling of specific astronomical phenomena. Additionally, the ability to process long texts (such as 128K tokens) will be strengthened to support automatic parsing and generation. At the same time, general capabilities will be maintained and optimized to ensure the model's stability and applicability across different domains. In the long term, general capabilities will continue to be optimized to maintain model stability and broaden applicability across different subfields of astronomy and related disciplines including astrochemistry, mathematical reasoning etc..

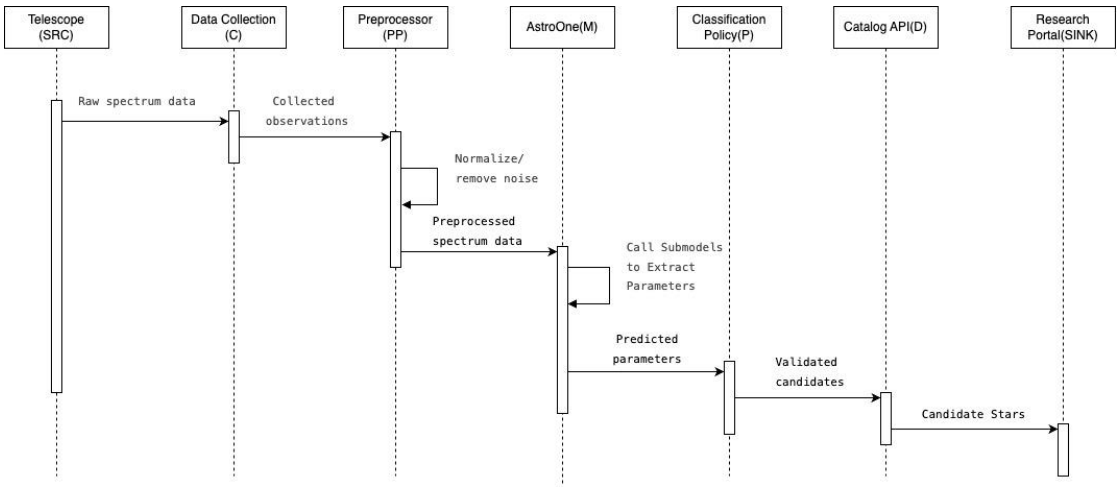
2. Data Collection and Augmentation: High-quality astronomical datasets will be expanded to include a wider variety of astronomical data. Data augmentation techniques will be applied to create more diverse training samples, addressing data scarcity and increasing both the diversity and breadth of the dataset. Long-term efforts will focus on curating specialized datasets aligned with emerging research challenges in astrophysics and cosmology.
3. Model Explainability and Scientific Validity: Explainable AI techniques will be introduced to make the decision-making process of the model more transparent, ensuring its scientific validity and credibility. Specifically, reinforcement learning (RL) will be explored to guide and expose the model's reasoning steps, providing clearer insight into how predictions are made. Regular error analysis will be conducted to identify and address the limitations of the model, ensuring its long-term stability and reliability.[5][12]
4. Internationalization and Open Collaboration: We plan to open-source models and datasets to promot global sharing and technological advancement. Short-term collaborations will focus on establishing partnerships with leading astronomical research institutions, AI research laboratories, and universities. In the long term, we plan to actively participate in international activities organized by ITU's Focus Group FG-AI4NDM (Focus Group on AI for Natural Disaster Management) and the IAU (International Astronomical Union), with the aim of fostering sustained multinational collaborations. By advancing astronomical research and strengthening international scientific cooperation, we strive to promote the well-being of humanity.

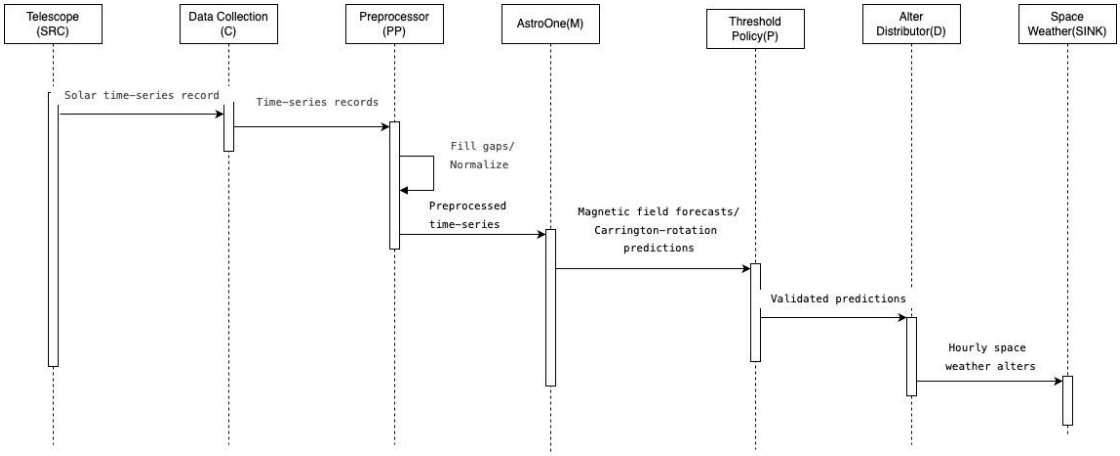
## 3 Use Case Requirements

You can access OneAstronomy by visiting https://OneAstronomy.zero2x.org/ through an internet-connected device. You can locally deploy OneAstronomy with 4 A100-80G GPUs.

- **REQ-01**: High-quality multimodal astronomical data for model training.
- **REQ-02**: Scalable cloud infrastructure for astronomical AI model deployment.
- **REQ-03**: Integration with existing astronomical databases and analysis tools.
- **REQ-04**: Real-time data analysis and predictive modeling capabilities for astronomical phenomena.

## 4 Sequence Diagram






## 5 References

[1] OneAstronomy. Accessed August 25, 2025. https://OneAstronomy.zero2x.org/.

[2] Liu, H., C. Li, Y. Lin, Z. Lin, P. Li, J. Zhou, and J. Tang. 2023. LLaVA: Large Language and Vision Assistant. arXiv preprint arXiv:2304.08485.

[3] Hinton, G., O. Vinyals, and J. Dean. 2015. Distilling the Knowledge in a Neural Network. arXiv preprint arXiv:1503.02531.

[4] Zhang, C., Y. Zhou, P. Li, and J. Tang. 2024. AstroLLaVA: An Instruction-tuned Large Multimodal Foundation Model for Astronomy. arXiv preprint arXiv:2402.01969.

[5] Liu, P., T. He, M. Chen, et al. 2023. Sci-Bench: Evaluating Scientific Reasoning and Knowledge in Foundation Models. arXiv preprint arXiv:2312.01463.

[6] Zhao, J., T. Wang, M. Yatskar, V. Ordonez, and K. Chang. 2017. "Men Also Like Shopping: Reducing Gender Bias Amplification Using Corpus-level Constraints." In Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing (EMNLP).

[7] Devlin, J., M. W. Chang, K. Lee, and K. Toutanova. 2019. "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding." In Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics (NAACL).

[8] Huang, Y. 2025. "SpecCLIP: Combining and Translating Spectroscopic Measurements for Stars." In Proceedings of IAU Symposium 397: UniversAI: Exploring the Universe with Artificial Intelligence.

[9] Tao, Y., and X. Zuo. 2025. "FALCO: A Foundation Model of Astronomical Light Curves for Time Domain Astronomy." In Proceedings of IAU Symposium 397: UniversAI: Exploring the Universe with Artificial Intelligence.

[10] Radford, A., J. W. Kim, C. Hallacy, A. Ramesh, G. Goh, S. Agarwal, S. Sastry, A. Askell, P. Mishkin, J. Clark, G. Krueger, and I. Sutskever. 2021. "Learning Transferable Visual Models from Natural Language Supervision." In Proceedings of the 38th International Conference on Machine Learning (ICML).

[11] Zhang, C., Y. Shen, Y. Zhou, and J. Tang. 2023. AstroCLIP: Contrastive Language-Image Pretraining for Astronomy. arXiv preprint arXiv:2311.15756.

[12] Yang, A., B. Yang, B. Zhang, et al. 2024. Qwen2.5 Technical Report. arXiv preprint arXiv:2412.15115.

[13] Weber, M., D. Fu, Q. Anthony, Y. Oren, S. Adams, A. Alexandrov, … and C. Zhang. 2024. RedPajama: An Open Dataset for Training Large Language Models. arXiv preprint arXiv:2411.12372.

[14] Soldaini, L., R. Kinney, A. Bhagia, D. Schwenk, D. Atkinson, R. Authur, … and H. Hajishirzi. 2024. Dolma: Data and Tools for Generating and Inspecting OLMo Pre-training Data. arXiv preprint arXiv:2402.00159.

[15] Hugging Face. "ShareGPT Vicuna Unfiltered." Accessed August 25, 2025. https://huggingface.co/datasets/anon8231489123/ShareGPT_Vicuna_unfiltered.

[16] Hugging Face. "MetaMathQA." Accessed August 25, 2025. https://huggingface.co/datasets/meta-math/MetaMathQA.

[17] Qwen. "Qwen2.5-72B-Instruct." Accessed August 25, 2025. https://huggingface.co/Qwen/Qwen2.5-72B-Instruct.

[18] OpenAI. "GPT-4o." Accessed August 25, 2025. https://platform.openai.com/docs/models/gpt-4o.

[19] de Haan, T., Y.-S. Ting, T. Ghosal, et al. 2025. AstroMLab 4: Benchmark-Topping Performance in Astronomy Q&A with a 70B-Parameter Domain-Specialized Reasoning Model. arXiv preprint arXiv:2505.17592.

[20] AstroMLab. "AstroSage-70B." Accessed August 25, 2025. https://huggingface.co/AstroMLab/AstroSage-70B.

[21] Hugging Face. "GPQA Dataset." Accessed August 25, 2025. https://huggingface.co/datasets/Idavidrein/gpqa.

[22] Wu, J., P. Liu, J. Deng, et al. 2023. InternVL-Chat: Vision-Language Foundation Model for Science with Scientific Multimodal Instruction Tuning. arXiv preprint arXiv:2312.01908.

# Use Case 8: AI-Based Signal Detection System for Residential TV Dish Alignment

**Organization**: Mbeya University of Science and Technology (MUST)

**Country:** Tanzania

**Contact Person(S)**:

Rehema Benson Richard, reykambona@gmail.com

sospeter j salawa, joseph.salawa@must.ac.tz

Cuthbert john Karawa, cuthbertjohnkarawa@gmail.com

## 1 Use Case Summary Table

| Item | Details |
|---|---|
| Category | Smart home/cities |
| Problem Addressed | Households frequently lose satellite signal due to manual misalignment or disturbances like wind. Additionally, one dish per service provider increases cost and clutter. |
| Key Aspects of Solution | Automated azimuth and elevation alignment using AI-powered feedback system. Supports access to multiple service providers. Automatically detects and corrects misalignment. |
| Technology Keywords | AI, Embedded Systems, Smart Dish, Signal Strength Feedback, Dish Automation |
| Data Availability | Private (manually collected from field technicians and simulation-based) |
| Metadata (Type of Data) | Annotations(Numerical (Azimuth & Elevation angles), signal strength, simulation logs) |
| Model Training and Fine-Tuning | Labelled data from manual installation used to calibrate alignment; further AI training planned post-deployment |
| Testbeds or Pilot Deployments | To be done as a prototype on university premises or residential environment |

## 2 Use Case Description

### 2.1 Description

Context and background

In many homes across Tanzania, residents face the inconvenience of losing satellite TV signal due to poor alignment of satellite dishes. These alignments are manually done and require high skill and time, and are often disturbed by weather conditions such as wind. The proposed project aims to solve this problem through an AI-based automated dish alignment system. This system enables the user to select a service provider (Azam, DSTV, Zuku), and then uses stored azimuth and elevation values to align the dish automatically via motor controls.

In case of disturbances that displace the dish, the system detects the change through signal strength feedback and re-initiates the realignment process to restore signal. The input mechanism uses a simple user interface to input desired service provider and angles if needed.

The expected outcome is a stable, low-cost, and accessible solution for automatic dish positioning, even in environments without technical expertise. It enhances convenience, saves time, and supports multiple providers using a single dish.

Objectives and Aims

- To design a cost-effective system that automates satellite dish alignment using AI logic.
- To enable one dish to access multiple service providers through auto-rotation.
- To ensure accurate alignment using reference nodes and feedback control.
- To reduce manual efforts and promote user-friendly interaction.
- To build a foundation for future simulations and prototype development.

Innovative Technological Approach

This project proposes a smart embedded system with the following features:

- An Arduino microcontroller that controls the physical rotation of the dish.
- A reference node that detects the current orientation of the dish.
- Planned integration of AI-based logic to detect signal quality and correct misalignment automatically.
- A simple user interface where users select the service provider, and the system aligns the dish accordingly.

Expected Impact and Benefits

- Enhanced signal reliability through automatic realignment.
- Reduced installation and maintenance costs by using a single dish for multiple providers.
- Increased user convenience by eliminating manual adjustment needs.
- Scalability for homes in underserved rural and urban areas.
- Support for digital learning and communication via stable TV reception.

**Use Case Status:** The use case is currently at the prototype development stage, with planned testing in university or residential environments.

## 2.2 Benefits of the use case

Industry, Innovation, and Infrastructure: The project promotes the development of reliable and sustainable communication infrastructure by utilizing smart systems to enhance signal distribution. By integrating intelligent satellite and set-top box coordination, the solution

represents a shift toward more efficient, software-defined communication networks that improve coverage and reduce signal interference.

Sustainable Cities and Communities: The solution supports digital inclusion by offering affordable access to television services, particularly in underserved or rural areas. This enhances access to information, education, and entertainment, fostering more connected and inclusive communities while reducing the digital divide.

Responsible Consumption and Production: The project minimizes equipment redundancy by enabling a single satellite dish to support multiple television service providers. This approach reduces the need for multiple installations, conserves material resources, and encourages more sustainable consumer behavior in the media and communication sector.

## 2.3 Future Work

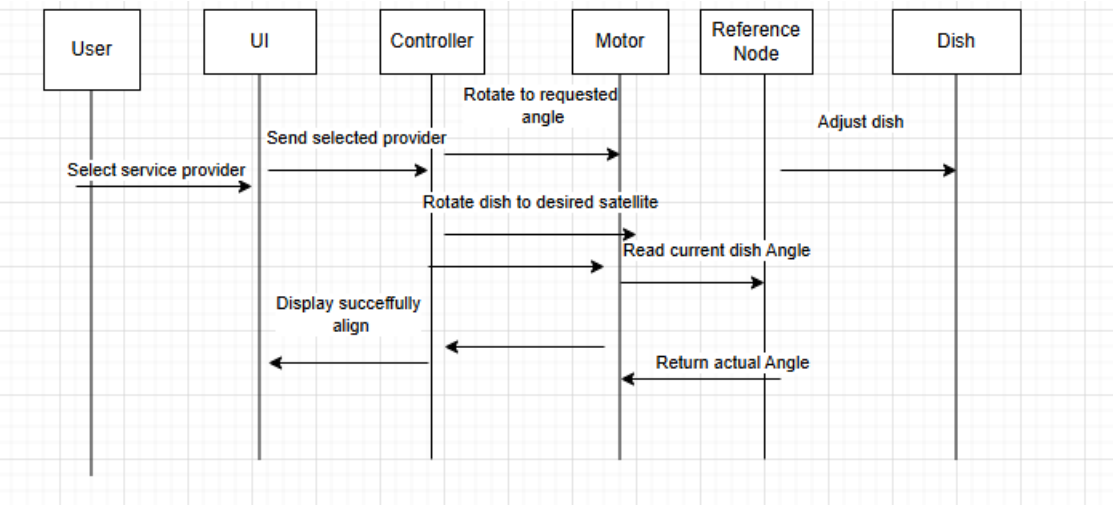
- Integrate IoT dashboard for remote control and updates.
- Implement real-time weather tracking to proactively adjust dish before disturbances.
- Train AI model with large dataset of misalignment patterns for faster realignment.
- Collaborate with dish service providers for pilot rollout.
- Optimize system size and reduce power consumption using solar-based designs.

# 3 Use Case Requirements

- **REQ-01**: Service Provider Selection and Retrieval of Pre-Stored Angle Data
  It is critical that the system allows users to select their desired satellite service provider (e.g., Azam, DSTV, Zuku) through a simple interface. Upon selection, the system should automatically retrieve and apply the corresponding pre-stored azimuth and elevation angle data for accurate dish alignment. This ensures a seamless user experience and supports multi-provider functionality without requiring technical expertise.
- **REQ-02**: Motorized Dish Alignment Based on Angle Input
  It is critical that the system includes a motorized mechanism capable of physically rotating the satellite dish in both azimuth and elevation directions based on the angle data input. This automated movement is essential for precise signal acquisition and eliminates the need for manual dish adjustments.
- **REQ-03**: Real-Time Feedback System for Signal Monitoring
  It is expected that a real-time feedback mechanism is integrated into the system to continuously monitor the signal strength received by the dish. This feedback loop enables the system to assess alignment accuracy and detect any performance degradation caused by misalignment or external disturbances.
- **REQ-04**: Automatic Realignment Capability After Disturbances
  It is critical that the system is capable of detecting misalignment events, such as those caused by wind or vibration, and then automatically initiating a realignment process using the stored reference data. This functionality ensures consistent signal availability and minimizes user intervention.
- **REQ-05**: Multi-Provider Switching Through Unified Interface
  It is of added value that the system provides users with the ability to switch between different service providers via a single, unified interface, with the dish automatically realigning based on the new provider's coordinates. This feature increases the utility of a single dish and offers flexibility to users who subscribe to multiple services.
- **REQ-06**: Continuous Signal Strength Monitoring and Logging
  It is expected that the system not only monitors signal strength during alignment but also

logs signal data over time to facilitate performance analysis and predictive maintenance. This functionality enhances reliability and supports further AI optimization based on signal trends.

## 4 Sequence Diagram




## 5 References

[1] Smith et al. (2018). "Automated Dish Alignment Systems Using Microcontrollers."

[2] Jones & Patel (2019). "Environmental Impact on Dish Signal Quality."

[3] Ahmed & Kumar (2020). "Real-Time Signal Processing for Satellite Dishes."

[4] Lee & Brown (2021). "Sensor-Based Feedback in Dish Tracking Systems."

[5] Zhang et al. (2022). "Machine Learning for Satellite Alignment Automation."

# Use Case 9: Smart Housing Experience

**Organization**: Mohammed Bin Rashid for Housing Establishment

**Country**: United Arab Emirates

**Contact Person(s)**:

Primary contact: Aseel Ayoub , Email: aayoub@mbrhe.ae

Secondary contact: Fatma Al ustad, Email: falustad@mbrhe.ae

## 1 Use case Summary Table

| Category | Smart home/cities |
|---|---|
| Problem Addressed | Accessing government housing services in Dubai used to be a long, complicated, and disconnected process. Customers often waited up to 7 years to complete their housing journey—from applying for a land grant to receiving their home. The system was burdened by manual approvals, duplicate paperwork, and poor coordination between key entities like banks, Dubai Municipality, and utility providers. This not only caused delays and confusion for citizens, but also increased operational costs and pressure on government employees. The lack of real-time data sharing and digital integration made it difficult to serve people efficiently or prioritize urgent social cases. |

(continued)

| Category | Smart home/cities |
|---|---|
| Key Aspects of Solution | 1. **AI-Powered Eligibility Engine**<br>Utilizes artificial intelligence to instantly evaluate customer applications, detect social urgency, and assign priority levels—reducing manual review time from months to seconds.<br>2. **End-to-End Service Automation**<br>Full digitization of the housing journey—from document submission to completion letter—reducing service time from 84 months to 32 months.<br>3. **Seamless API Integration with Government & Banking Entities**<br>Real-time data exchange with Dubai Land Department, DEWA, Dubai Municipality, and Emirates Islamic Bank for instant approvals and status updates.<br>4. **Pre-Approved Smart Housing Designs**<br>A library of digitally approved architectural plans speeds up engineering approvals, reducing design approval time by 85% and saving over AED 29 million in costs.<br>5. **Customer-Centric Digital Dashboard**<br>An integrated platform where customers can track their application progress, receive alerts, and communicate with service agents—enhancing transparency and satisfaction.<br>6. **Document Digitalization & AI-Powered Pipeline**<br>The service pipeline includes digital document ingestion, automated evaluation through AI models, and real-time interface modules with citizens via dashboard and notifications. |
| Technology Keywords | AI, RPA (Robotic Process Automation), predictive analytics, system integration, APIs, e-government, data interoperability, digital identity |
| Data Availability | Public / Private (MBRHE internal data, stakeholder APIs, anonymized customer feedback datasets) |
| Metadata (Type of Data) | Text (forms, documents), Visual (design templates), Structured (application and eligibility data), Temporal (timelines, service stages) |
| Model Training and Fine-Tuning | AI models were trained using MBRHE historical customer applications, feedback logs, and partner data. Fine-tuning involved supervised learning for eligibility scoring, predictive modeling for urgency detection, and integration optimization for APIs. |
| Testbeds or Pilot Deployments | Pilot phase completed with select customers and partners in 2023; full deployment started in 2024. Stakeholder feedback and usage data continuously enhance the model. |
| Code Repositories | Not publicly available; proprietary solution developed in collaboration with private tech vendors. |

## 2 Use Case Description

### 2.1 Description

The Smart Housing Experience is an end-to-end AI-powered digital transformation initiative led by the Mohammed Bin Rashid Housing Establishment (MBRHE) in Dubai. The initiative addresses the challenge of fragmented and delayed housing services by offering a unified,

intelligent, and predictive system that connects beneficiaries with land grants, loans, construction, maintenance, and support services in one seamless journey.

The pipeline begins with document digitalization, followed by automated AI evaluation, and concludes with citizen interface modules through a digital dashboard and mobile services. This modular design enables efficient data intake, real-time eligibility assessment, and seamless interaction with stakeholders and customers.

### Main Features:

- Real-time AI-driven service recommendations
- Automated loan and grant approvals within minutes
- Pre-approved housing designs to speed up construction
- End-to-end API integration with Dubai Land Department, DEWA, Dubai Municipality, Emirates Islamic Bank, and others
- Customer dashboard for service progress, alerts, and communication

### AI Models Trained:

- Eligibility Scoring Models: Classify customers based on financial, social, and historical data
- Urgency Detection Models: Identify priority social cases using sentiment and profile analytics
- Predictive Analytics Models: Recommend housing solutions based on service availability and forecast demand
- Robotic Process Automation (RPA): Automate repetitive tasks and document processing

### Expected Benefits:

- Reduced service timeline from 84 months to 32 months
- 10x reduction in service requirements
- AED 29M saved for customers via pre-approved housing designs
- AED 246M in annual operational cost savings
- 183% improvement in productivity and throughput
- Customer satisfaction increased to 98.75%

Use Case Status:

### Partners

**MBRHE (Project Owner)**: Led strategy, data access, and platform development

**Emirates Islamic Bank**: Housing loan disbursement and service integration

**Dubai Land Department**: Property registration and title transfer integration

**Dubai Municipality**: Construction permit and land zoning approvals

**DEWA**: Utility setup and connection automation

**Private Tech Vendors**: AI model development and RPA implementation

**Customers**: Participated in co-design sessions, surveys, and testing

## 2.2 Benefits of use case

- By streamlining access to affordable, sustainable housing solutions through smart technology.
- Through AI-driven service delivery, integrated platforms, and digitized infrastructure management.
- Enhancing transparency, reducing corruption risks, and improving citizen trust in government processes.
- Supporting vulnerable families with faster access to essential housing services and financial support.

## 2.3 Future Work

### Predictive Analytics for Proactive Housing Services

- Use AI to predict customer needs based on lifestyle data and housing trends.
- Offer services before customers request them (e.g., maintenance reminders, loan upgrades).

### Hyper-Personalization of Housing Recommendations

- Use behavioral data and AI to suggest tailored housing options based on family size, future planning, financial behavior, and life events.

### Blockchain for Secure Document Exchange

- Explore blockchain-based smart contracts for agreements with banks, contractors, and customers to enhance trust and eliminate paper processes.

# 3 Use Case Requirements

- **REQ1**: Secure access to MBRHE customer and service history data
- **REQ2**: Full integration with stakeholder APIs (DLD, Municipality, DEWA, Bank)
- **REQ3**: AI engine for eligibility checks
- **REQ4**: Data protection and governance framework aligned with Dubai Digital Authority
- **REQ5**: Scalable cloud infrastructure for high-volume processing
- **REQ6**: User-friendly front-end with multilingual and accessibility support
- **REQ7**: Real-time dashboard for service monitoring and reporting
- **REQ8**: Unified and secure access to services via UAE PASS

## 4 Sequence Diagram

## 5 References

[1] Government of Dubai. Dubai Social Agenda 2033: Aligning with Long-Term National Goals. Accessed June 19, 2025.

[2] Government of Dubai. Dubai Digital Strategy: Supporting Digital Integration, AI, and UAE PASS Use. Accessed June 19, 2025.

[3] Government of Dubai. Dubai Paperless Strategy: Advancing Digital Document Workflows and Automation. Accessed June 19, 2025.

[4] Government of Dubai. Dubai Government Customer Happiness Index: Benchmarking MBRHE's No. 1 Ranking in Customer Satisfaction. Accessed June 19, 2025.

# Use Case 10: AI-Powered Network Quality Prediction for Seamless Online Event Planning

**Organization**: AI4Africa

**Country**: Nigeria

**Contact Person(s)**:

Emmanuel Aaron (aaronemmanuel054@gmail.com, +234807200689),

Dr. Houda Chihi (houda.chihi@supcom.tn),

Blessed Guda (gudablessed@gmail.com),

Emmanuella Sule (suleemmanuella0010@gmail.com),

Chidi Ebube (chidizack24@gmail.com),

Emmanuel Ani (ani.mlengineer@outlook.com),

Okafor Miracle Uche (okaformiracle212@gmail.com)

## 1 Use Case Summary Table

| Item | Details |
|---|---|
| Category | Smart home/cities |
| Problem Addressed | In developing regions such as Africa, frequent and unpredictable network disruptions lead to missed opportunities for students, job seekers, and professionals. Candidates end up losing online interviews from sudden connection losses, while students are disrupted during online examinations. Businesses and virtual conferences are also disrupted, leading to lost time and revenue. |
| Key Aspects of Solution | Federated Generative Neural Networks & LSTMs to predict optimal period of network functionality |
| Technology Keywords | Network Failure Prediction |
| Data Availability | some other private data sources could be from collaboration with ISPs, there are also open source network simulation platforms such as, NS-3, Mini net, OMNET++, CORE, GNS3[1] |
| Metadata (Type of Data) | Numerical data (ping tests, upload speed, download speed, Number of active devices in region), textual data (weather parameters e.t.c) |
| Model Training and Fine-Tuning | For Predictive Modelling, the system will leverage Federated Graph Neural Networks (FedGNN) alongside Long Short-Term Memory (LSTM) and Transformer based models |

(continued)

| Item | Details |
|---|---|
| Testbeds or Pilot Deployments | Anmol Gupta, "Internet Speed," *Kaggle Datasets*. Accessed June 19, 2025. [3] |
| Code repositories | N/A |

# 2 Use Case Description

## 2.1 Description

Stable internet connections are necessary in order to partake in online meetings, work interviews, internet courses, and online conferences. However, in less developed regions like Africa, the constant and predominantly unforeseeable network crashes mean opportunities go by for pupils, job applicants, and working professionals. Job interviewees miss valuable interviews due to abrupt shutdowns, and pupils are interrupted midway through imperative online exams. This disruption influences companies and virtual gatherings as well, leading to missed productivity and revenues.

While there are some present mechanisms that might detect or even anticipate network failures, their prime objective is to inform the service providers or carry out retrospective analysis. This application, however, proposes an AI-based network quality prediction system which not only predicts failures but also assists users in planning ahead for their online behaviour.

The system leverages a combination of past network history, real-time user feedback, weather conditions, and telecom operator feeds to provide accurate, location-specific predictions of network quality. Users can input a time and place to receive AI-based predictions of network stability. Crucially, the platform goes beyond basic passive notifications by offering recommendations on the optimal available time slots for online events, enabling users to schedule with confidence.

It seamlessly integrates with popular applications such as Google Calendar, Zoom, and Microsoft Teams, giving reminder notifications to reschedule or organize contingency options in advance. This anticipation foresees and mitigates risks of network failures and provides more efficient online experiences.

Unlike typical telecom coverage maps providing a general availability or speed test applications that provide one-off measurements, this system provides dynamic, forward-looking information. It shifts the paradigm from reactive to proactive management of connectivity. This kind of proactive functionality is especially helpful in low-connectivity environments, where digital inclusion is riding on the power to plan effectively beyond infrastructure limitations.

By maximizing the utilization of network resources in that and at which times they are accessed, the system maximizes overall digital efficiency and allows users to avoid disruptions before they occur. One discussed problem is that some data sets—particularly those that come from telecom firms—may not be readily accessible for training, but federated data collection and crowdsourcing methods eliminate this issue.

With real-time and forecast network quality data, the solution not only aims to detect failures but to enable better digital planning, reduce lost opportunities, and enhance online experiences, especially in areas prone to network instability.

**Use Case Status: N/A**

## 2.2 Benefits of use case

By leveraging AI-driven network optimization, this solution contributes to the development of a more robust and resilient digital infrastructure. In many developing economies, including Africa, incomplete network coverage results in unreliable connectivity. AI-powered predictive analytics improve the reliability and efficiency of telecommunication infrastructure, reducing service interruptions in remote work, online education, and e-governance. Improved network reliability fosters economic growth by enabling online enterprises, remote work opportunities, and digital innovation hubs.

Ensuring reliable internet connectivity in both urban and rural areas is key to building sustainable digital societies. Urban environments rely heavily on stable internet for business operations, public services, and education, while rural communities face even greater challenges in accessing quality network services. By providing predictive insights, this AI-driven platform empowers residents, businesses, and institutions to plan effectively, mitigating the impact of poor connectivity on daily life. Enhanced digital access also supports smart city initiatives and fosters technology-inclusive development.

Collaboration with telecom providers will refine data collection and improve forecasting accuracy. Integrating predictive analytics into network data collection will enhance forecasting precision. AI researchers will contribute their expertise in developing robust predictive models, while policymakers can leverage insights from this system to strategize for improved digital infrastructure and increased digital inclusion. By fostering multi-stakeholder collaboration, this use case will drive collective action toward better internet accessibility, bridging the digital divide, and supporting sustainable technological advancements.

## 2.3 Future Work

With scholarships and funding, the way forward for this project will be to further develop data collection, build a proof of concept, train AI models, integrate with existing platforms, and normalize the framework for adoption in general.

The first phase involves aggregating diverse datasets through partnerships with telecom providers to access real-time network logs, historical performance metrics, and Quality of Service (QoS) indicators. To address data accessibility limitations, federated data collection will be implemented via crowdsourced user reports. Users will voluntarily submit network performance data, which, when combined with real-time weather conditions and telecom infrastructure updates, will enhance predictive accuracy. Graph-based representations of network topology will be constructed, mapping base stations, traffic congestion points, and interference zones for structured learning.

For predictive modeling, the system will leverage Federated Graph Neural Networks (FedGNN) alongside Long Short-Term Memory (LSTM) networks and Transformer-based models.

FedGNN will model spatiotemporal dependencies within the network graph, learning relationships between towers, nodes, and connectivity fluctuations without centralizing sensitive telecom data. This decentralized approach ensures privacy while allowing efficient learning across distributed data sources.

LSTM networks will handle sequential dependencies in time-series forecasting, learning from past trends to predict future network performance.

Transformer models, particularly Temporal Graph Transformers, will enhance long-range dependency learning, refining predictive insights across multiple timescales.

Telecom API and scheduling tool integration will be a key milestone. The platform will incorporate network prediction intelligence into widely used applications like Google Calendar, Zoom, and Microsoft Teams to enable users to receive notifications when scheduling online events. An API will also be developed to enable mobile applications and enterprise solutions.

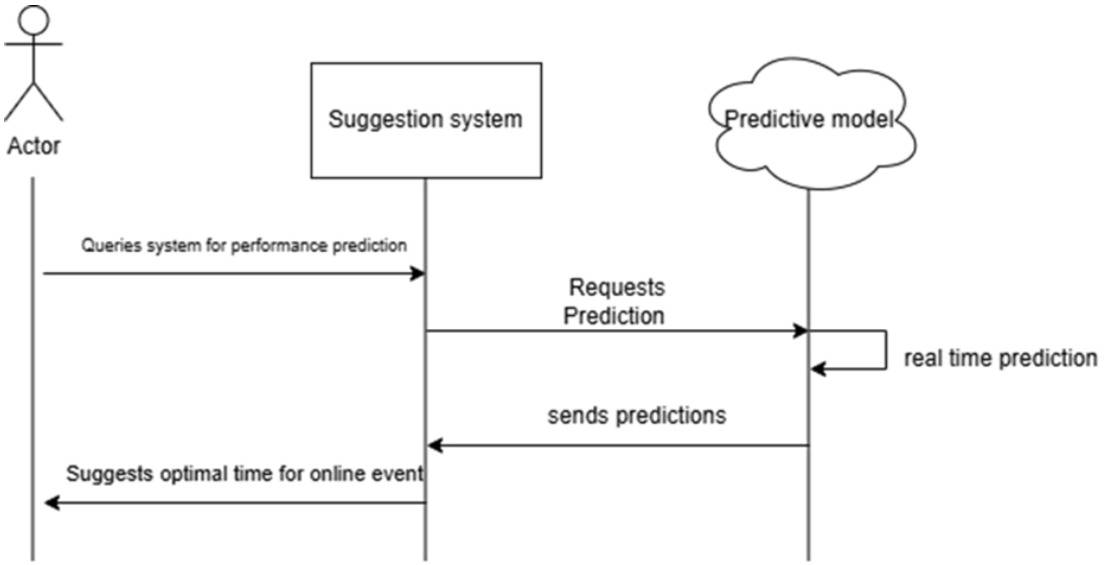
To ensure standardization, the project will collaborate with ITU and telecom regulatory authorities to develop a framework for AI-powered network prediction. Open-sourcing the model will ensure transparency as researchers and developers can continuously improve it.

Finally, real-world and simulation testing will be conducted in various locations, including busy city centers and rural towns. Pilot tests will ascertain the model's effectiveness, measure its impact on network provider efficiency, and quantify user experience improvements

## 3 Use Case Requirements

- **REQ-01**: The system shall collect historical and real-time network data from telecom providers, weather APIs, and user feedback to support accurate predictions.
- **REQ-02**: The system shall implement Federated Graph Neural Networks (FedGNN) to model network behavior while preserving data privacy.
- **REQ-03**: The platform shall detect anomalies and provide real-time alerts on potential network disruptions to prevent online session failures.
- **REQ-04**: The system shall integrate seamlessly with tools such as Google Calendar, Zoom, and Microsoft Teams for automatic scheduling support and notifications.
- **REQ-05**: Users shall be able to input a specific location and time to receive AI-powered predictions of network quality.
- **REQ-06**: The system shall recommend optimal time slots for scheduling online events based on historical network trends and forecasted conditions.
- **REQ-07**: The platform shall allow users to submit feedback on actual network performance to continuously improve prediction accuracy.

## 4 Sequence Diagram




## 5 References

[1] nternational Telecommunication Union. "The Future of Fixed Broadband." *ITU Journal: ICT Discoveries* 4, no. 3 (2023): Article A31. Accessed June 19, 2025. https://www.itu.int/dms_pub/itu-s/opb/jnl/S-JNL-VOL4.ISSUE3-2023-A31-PDF-E.pdf.

[2] Michele Polese, Marco Giordani, Michele Zorzi, and Douglas Castor, "A Survey on 5G Millimeter Wave Communications for UAV-Assisted Wireless Networks," *IEEE Access* 6 (2018): 61070–61093. Accessed June 19, 2025. https://ieeexplore.ieee.org/document/8407833/.

[3] Anmol Gupta, "Internet Speed," *Kaggle Datasets*. Accessed June 19, 2025. https://www.kaggle.com/datasets/getanmolgupta01/internet-speed.

# Use Case 11: AI-Powered Space Optimization for Low-Income Housing

**Organization:** VIT University

**Country**: India

**Contact Person(s)**:

Mithun Balaji V – mithunbalaji.v2022@vitstudent.ac.in mailto:mithunbalaji.v2022@vitstudent.ac.in

Sankar Raja S - sankar.raja2022@vitstudent.ac.in

Mentor: Dr. Geetha S - geetha.s@vit.ac.in

## 1 Use Case Summary Table

| Item | Details |
|---|---|
| Category | Smart home/cities |
| Problem Addressed | Inefficient use of space in government-built low-income housing |
| Key Aspects of Solution | AI-driven layout optimization, RL-based adaptive planning, constraint-based filtering, and CV for layout analysis |
| Technology Keywords | Deep Learning, Reinforcement Learning, Constraint Solvers, Floor Plan Generation, Computer Vision, 2D/3D Spatial Analysis |
| Data Availability | Public (ZInD, FloorPlanCAD, HouseExpo, SYNBUILD-3D, UN-Habitat, HUD) |
| Metadata (Type of Data) | 2D CAD floor plans, 3D building wireframes, image data (PNG/JPEG), spatial graphs, furniture dimensions, user preferences |
| Model Training and Fine-Tuning | CNNs for layout analysis, RL (PPO/DQN) for layout optimization, constraint filtering with OR-Tools |
| Testbeds or Pilot Deployments | Concept tests using FloorPlanCAD and HouseExpo layouts; pilot simulations planned with Indian low-income floor plans |
| Code repositories | [1] |

# 2 Use Case Description

## 2.1 Description

### Context and Background:

Government housing schemes in India aim to provide shelter for economically weaker sections, but often the layouts are poorly optimized. These homes become difficult to navigate, overcrowded, and underutilized. Traditional space planning is manual, static, and lacks customization based on user needs.

### Objective and Aim:

This use case proposes a smart AI system that suggests optimal furniture and storage layouts within constrained room dimensions using reinforcement learning, deep learning, and rule-based filtering. The goal is to improve the livability of small homes without increasing cost or requiring redesign.

### Technological Approach:

### 2D/3D Layout Encoding:

- CAD plans are parsed into 2D grid encodings.
- 3D voxel maps simulate volumetric constraints.
- Graph-based room representations encode furniture relationships.

### Computer Vision + CNNs:

- Analyze spatial inefficiencies and extract layout patterns from real floor plans.

### Reinforcement Learning:

- An agent (PPO/DQN) optimizes layout decisions by simulating placement actions and receiving reward feedback based on space usage, accessibility, and aesthetics.

### Constraint Filtering:

- OR-Tools filters infeasible layouts (e.g., blocked pathways, overlapping items) to ensure usability.

### Expected Impact:

- Enhanced user comfort and usability in compact homes
- Efficient use of taxpayer-funded government housing
- Scalable optimization adaptable to various living arrangements (hostels, dorms, rental units)

### Use Case Status:

Currently in prototyping phase. Synthetic data from FloorPlanCAD and SYNBUILD-3D used for training and evaluation. Real-world layout data integration in progress.

Partners

**Academic:** VIT University (Model development, simulation, evaluation)

**Mentor:** Dr. Geetha S – Technical oversight and guidance

**Planned Collaborations:** NGOs in Tamil Nadu & housing development boards for pilot deployments and feedback validation

## 2.2 Benefits of use case

By enhancing space efficiency in government-built housing, this use case promotes equity in urban design and contributes to building smarter, more resilient communities.

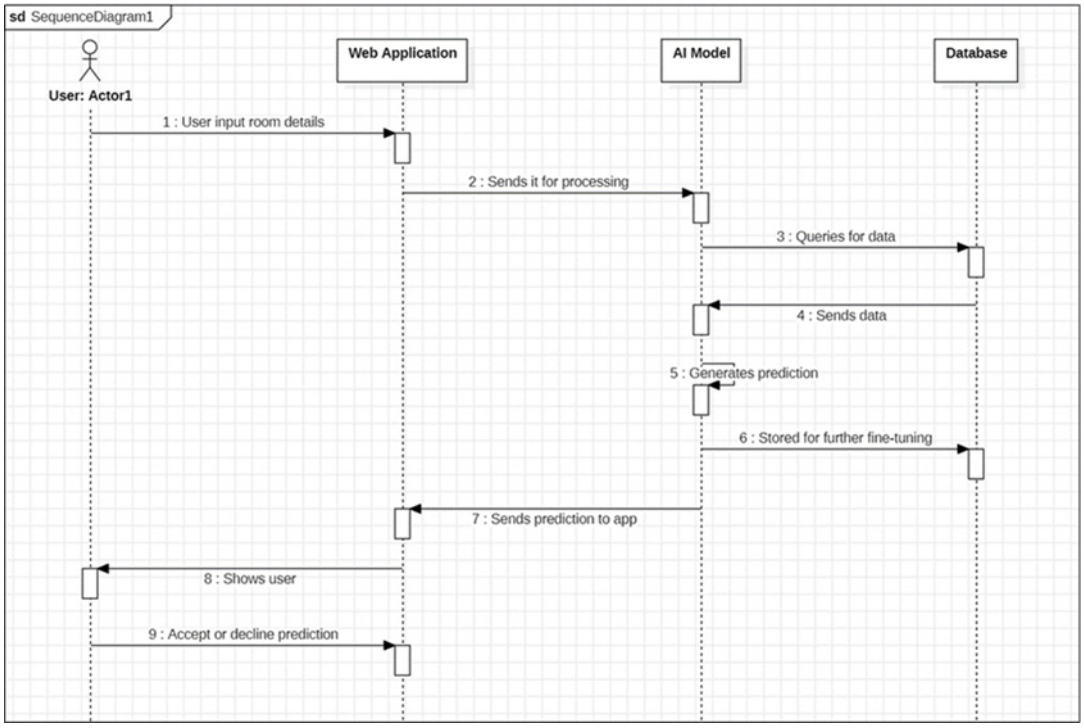
## 2.3 Future Work

- Integration of AR/VR for interactive visualization
- Real-world pilot deployment with housing boards
- RL model fine-tuning using user feedback
- Expansion to include smart appliances, modular furniture, and rental layouts

# 3 Use Case Requirements

- **REQ-01**: 2D/3D layout datasets with annotated dimensions and semantics
- **REQ-02**: Furniture metadata repository (sizes, orientations, use cases)
- **REQ-03**: Deep Learning pipeline for layout extraction and CNN-based pattern recognition
- **REQ-04**: RL policy network (e.g., PPO) for layout decision-making
- **REQ-05**: Constraint solver (OR-Tools) to ensure practical, walkable designs
- **REQ-06**: Visualization module (matplotlib / WebGL) for user output
- **REQ-07**: User interface for manual input and preference settings
- **REQ-08**: Modular AI backend for real-time recommendation generation

## 4 Sequence Diagram




## 5 References


[1] Di, X., & Yu, P. Q. (2021). Deep Reinforcement Learning for Producing Furniture Layout in Indoor Scenes. *arXiv preprint arXiv:2101.07462*. https://arxiv.org/abs/2101.074621

[2] Roy, U. K., & Roy, M. (n.d.). Space standardisation of low-income housing units in India. *Channel i*. Retrieved from https://channeli.in/media/faculty_biodata/papers/caf820fa-5f9d-4774-8e80-cb997b8a22ab.pdf2

[3] Wang, Y., et al. (2024). I-Design: Personalized LLM Interior Designer. *arXiv preprint arXiv:2404.02838*. https://arxiv.org/html/2404.02838v13

[4] Apple Inc. (2022). Create parametric 3D room scans with RoomPlan. *WWDC22*. https://developer.apple.com/videos/play/wwdc2022/10127/4

[5] Wang, X., et al. (2023). Floor Segmentation Approach Using FCM and CNN. *Applied Technology and Artificial Intelligence in Machine Learning*, 2(1). https://www.acadlore.com/article/ATAIML/2023_2_1/ataiml0201045

[6] van Rees, E. (2021). RIEGL 3D Voxel Analysis Module enables surface comparisons in underground, cave, or tunnel environments. *Geo Week News*. https://www.geoweeknews.com/news/riegl-3d-voxel-analysis-module-enables-surface-comparisons-in-underground-cave-or-tunnel-environments6

[7] Planner 5D. (2024). AI floor plan generator. https://planner5d.com/use/ai-floor-plan-generator7

[8] Voxelgrid. (2024). 2D digitalization. https://www.voxelgrid.com/en/2d-digitalization?tab=09

[9] Maket AI. (2024). Maximizing Efficiency in Architecture with AI Floor Plan Generators. https://www.maket.ai/post/maximizing-efficiency-in-architecture-with-ai-floor-plan-generators10

[10] Di, X., & Yu, P. Q. (2021). Deep Reinforcement Learning for Furniture Layout Simulation in Indoor Graphics Scenes. *SIMDL*. https://simdl.github.io/files/2.pdf

# 4.9. Category 9: Accessibility

## Use Case 1: Bridging the Digital Gap: AI Sign Language Interpretation Solution

**Organization**: Baidu

**Country**: China

**Contact Person(s)**:

Primary: Tianmin Long, intlcomm@baidu.com

Secondary: Yang Liu, liuyang145@baidu.com

Additional Contact: Cathy Fu, fuduanling@baidu.com

### 1 Use Case Summary Table

| Item | Details |
|---|---|
| Category | Accessibility |
| Problem Addressed | Public accessibility issue for the deaf and hard-of-hearing community |
| Key Aspects of Solution | Artificial Intelligence(AI) Sign Language Platform, AI Sign Language All-in-One Machine, and digital avatar |
| Technology Keywords | AI, sign language translation and interpretation, digital avatars |
| Data Availability | Private |
| Metadata (Type of Data) | Text, audio, visual<br>(Note: Data were collected through a dedicated collection app and are strictly limited to model training purposes only.) |
| Model Training and Fine-Tuning | The platform has collected nearly 10 million natural sign language samples to build a large-scale sign language corpus to generate natural sign language that aligns with the deaf and hard-of-hearing(DHH) community's communication habits. |
| Testbeds or Pilot Deployments | University and Research Lab Pilot, Rehabilitation Center Integration, Government Service Counter Trials [1] |

# 2 Use Case Description

## 2.1 Description

For more than 27.8 million DHH individuals in China, sign language interpretation is an essential service. However, the number of certificated sign language interpreters in China was less than 500 in 2023, representing a huge talent gap in the sign language interpretation market.

Meanwhile, developing a sign language translation system is uniquely challenging. Sign language differs fundamentally from spoken Chinese in grammar, structure, and expression, requiring key information extraction rather than direct translation. As a low-resource language, it also lacks sufficient data for AI model training. Moreover, natural sign language communication relies not only on hand gestures but also on mouth movements and body language, demanding complex multi-channel synthesis.

To address this issue, Baidu AI Cloud's XiLing digital avatar platform launched solutions including the Baidu AI Sign Language Platform[1], AI Sign Language All-in-One Machine, and digital avatar, aiming to empower the DHH community to integrate into social life and equally benefit from technological advancements. Unlike earlier solutions that could only translate speech into sign language, Baidu AI sign language interpretation tools realize the bidirectional translation between sign language and text, speech and regular videos, enabling truly natural two-way communication.

AI Sign Language Platform provides on-demand automated sign language translation and interpretation through digital avatars [2], while ensuring consistency and accuracy. The Platform supports typical sign language translation scenarios such as bilingual video production and sign language livestream. It integrates AI capabilities like automatic speech recognition (ASR) and sign language translation with a modular business platform and a digital avatar rendering engine[3], enabling efficient data flow from audio/video/text input to final media output via a unified pipeline.

The use case addresses these challenges with a three-step approach: converting speech to text, translating text into sign language codes, and generating expressive sign language via digital avatars, thanks to its robust suite of AI capabilities.

Additionally, the device is trained on a dataset comprising nearly ten million "natural sign language samples," ensuring its ability to provide accurate and nuanced sign language interpretation. It can achieve a near-field speech recognition rate on mobile phones of over 98%. The AI sign language virtual interpreter also has 98.5% accuracy of mouth shape generation.

AI Sign Language All-in-One Machine is the latest offering in XiLing's comprehensive suite of solutions. With a creative and interactive design of dual screens, the Machine achieves two-way real-time translation between sign language and text. As a typical application integrating multimodal AI capabilities, especially in visual recognition and natural language processing, the Machine tackles challenges such as accurately tracking and recognizing sign language against an unsharp background and fast motion of sign language movements.

To address the common problem of low recognition rates for long sentences in attention models, XiLing has uniquely innovated with the introduction of the world's first online attention model for voice interaction. This model's streaming attention capabilities not only outperform

those of whole-sentence attention but also offer stable solutions to noise, effectively mitigating latency issues in online services.

XiLing's end-to-end speech recognition model directly maps speech to text, which has resulted in over a 15% improvement in cloud-based recognition accuracy. The adoption of unsupervised learning is set to become a standard in industrial applications, leading to a marked increase in both whole-sentence and dialect recognition rates. This advancement enhances communication accessibility and accuracy for a wider audience, marking a significant leap forward in the field of speech recognition and sign language interpretation.

Its scalability allows deployment across a wide range of platforms and scenarios. It supports online platform deployment within hours or offline all-in-one machine plug-and-play, as well as integration into various digital channels such as TVs, apps, websites, and WeChat mini-programs, offering "barrier-free deployment" for the DHH community to access sign language services in their daily lives. The platform's cost-effectiveness makes sign language services more affordable, reaching a wider audience than human interpreters.

**Use Case Status:** During the 2022 Winter Olympics, the XiLing-powered AI sign language anchors provided 24/7 sign language live broadcasting and play-by-play for the DHH community, garnering over 100 million views, including both live broadcasts and video replays[4]

**Partners: Tianjin University of Technology School for the Deaf, language linguists, and special education experts**

## 2.2 Benefits of the use case

Good Health and Well-being: AI-powered sign-language solution employs the Streaming Multi-layer Truncated Attention Model (SMLTA) for online automatic speech recognition (ASR) and neural-network translation—refined with linguistic expertise—to achieve over 98% speech recognition accuracy and 98.5% sign-language fidelity. By deploying AI interpreters in hospitals and telehealth platforms, it removes barriers that 97% of deaf patients face during medical consultations, enabling round-the-clock access to healthcare information and services.

Quality Education: Adaptive Natural Language Processing(NLP) algorithms tailor sign-language outputs to regional dialects and classroom contexts. DHH students gain real-time 3D avatar translations of STEM (Science, Technology, Engineering, Mathematics) lectures, breaking down language barriers and ensuring that deaf learners receive the same depth of instruction as their hearing peers. This promotes truly inclusive and equitable learning environments.

Decent Work and Economic Growth: By integrating AI interpreters into workplaces and customer-facing services—such as banks and government offices—the platform dismantles communication obstacles that previously limited employment and entrepreneurship opportunities for deaf individuals[5]. Organizations report smoother workflows and broader talent pools as hearing-impaired staff participate fully in meetings, training, and client interactions.

Reduced Inequality: The solution standardizes sign-language support across urban centers and remote regions by training on a corpus of over 11,000 gestures and regional variations. Whether in legal aid clinics or municipal announcements, deaf users receive uniform, high-quality interpretation—narrowing the service gap between city and countryside.

Sustainable Cities and Communities: The plug-and-play All-in-One Machine embeds accessibility into everyday urban infrastructure—airports, hospitals, banks, and public transit hubs—while 4D-scanned digital avatars (98.5% mouth-shape accuracy) stream sign-language feeds to TVs, apps, and kiosks. Proven at the 2022 Winter Olympics, this scalable deployment ensures that barrier-free communication becomes a lasting feature of smart, inclusive cities.

## 2.3 Future Work

The planned enhancement of this solution will focus on the following domains:

Data Collection: We will expand our training datasets to include a wider range of sign language variations and dialects. This will involve collaborating with local communities, sign language experts, and organizations serving the DHH community to ensure our data is comprehensive and inclusive.

Model Development: We will continue to refine and enhance the platform's capabilities, focusing on improving the accuracy and richness of sign language vocabulary, motion capture technology, and lip shape correction. Our goal is to upgrade our current models to enhance performance, reduce latency, and add real-time conversational abilities.

Privacy Enhancement: We will explore the implementation of federated learning techniques to improve privacy during on-device model updates. This will involve developing methods that allow models to learn from decentralized data sources without the need for raw data to be shared, thus enhancing user privacy.

Performance Optimization: We will focus on developing and integrating lightweight transformer models into our system. These models will be designed to optimize performance across a variety of hardware platforms, ensuring that our sign language interpretation services are accessible and efficient on different devices.

Create New Variations/Extensions to the Same Use Case: We will extend the deployment of these AI sign language interpretation solutions to cover a broader range of public services, including hospitals, banking systems, transportation hubs, airport facilities, commercial complexes, educational institutions, and municipal infrastructure. This expansion strategy focuses on adapting our proven technical frameworks to diverse environmental requirements while maintaining core accessibility principles, ultimately creating an inclusive ecosystem that elevates barrier-free services across urban living spaces.

# 3 Use Case Requirements

- **REQ-01**: It is critical that the platform employs advanced visual recognition techniques that can accurately track and interpret fast hand movements without losing detail. This includes leveraging high-speed cameras or enhanced image processing algorithms to reduce motion blur and ensure precise gesture recognition.
- **REQ-02**: It is critical that the system incorporates robust background noise filtering and object detection capabilities to isolate hand gestures from noisy or cluttered environments. This may involve using advanced computer vision models to focus on the hands and gestures even in visually complex settings.
- **REQ-03**: It is critical that the platform uses diverse training datasets that include contextual sign language usage, not just isolated signs. This ensures the model can accurately

interpret signs in real-world conversations and reduce the risk of misinterpretation due to lack of contextual understanding.

- **REQ-04**: It is critical that sign language users are kept involved and contributing to the production, processing, and annotation of training datasets, ensuring that the training data is both authentic and reflective of the varied sign language usage in real-world scenarios. This ensures the model's accuracy in capturing the nuances of sign language.
- **REQ-05**: It is critical that the solution accounts for regional differences in sign language, addressing variations in signs and dialects across different areas to ensure widespread applicability and inclusivity.

## 4 Sequence Diagram

## 5 References

[1] Baidu. (2023). Baidu AI Sign Language Platform [Software platform]. https://persona.baidu.com:8891/csl

[2] Baidu, Inc. (2022). Sign language translation method, device, equipment, and computer-readable medium (Patent No. CN115376487A). China National Intellectual Property Administration. https://patents.google.com/patent/CN115376487A/en

[3] Baidu, Inc. (2024). Sign language recognition method and apparatus, electronic device, and storage medium (Patent No. CN117541696B). China National Intellectual Property Administration. https://patents.google.com/patent/CN117541696B/en

[4] Ling, T. (2021, December 2). Chinese TV introducing AI sign language presenter at the next Olympics. Metro. https://metro.co.uk/2021/12/02/chinese-tv-introducing-ai-sign-language-presenter-at-the-next-olympics-15700396/

[5] Mozur, P. (2022, July 21). A digital human could be your next favorite celebrity—or financial advisor. MIT Technology Review. https://www.technologyreview.com/2022/07/21/1056335/a-digital-human-could-be-your-next-favorite-celebrity-or-financial-advisor/

# Use Case 2: AI-Driven Drowsiness Detection: Enhancing Road Safety through Inclusive and Real-Time Monitoring

**Organization**: Catalink Limited

**Country**: Cyprus

**Contact Person(s)**:

Konstantinos Avgerinakis, koafgeri@catalink.eu

Zenon Lamprou, zenon.lambrou@catalink.eu

Maria Eleni Skarkala, maria.skarkala@catalink.eu

## 1 Use Case Summary Table

| Item | Details |
|---|---|
| **Category** | Accessibility |
| **Problem Addressed** | Current Driver Monitoring Systems (DMS) struggle to accurately detect drowsiness leading to false negatives and reduced reliability due to biases in the models. These issues stem from insufficient diversity in training data and overly generalized models, resulting in biased performance. To ensure fairness, accuracy, and inclusivity, it is essential to address these limitations and improve the ethical deployment of AI in driver monitoring systems. |
| **Key Aspects of Solution** | Training unbiased and fair models to be able to detect drowsiness detection when driving. This solution is primarily designed for public transport, but it can also be utilized by individual drivers. |
| **Technology Keywords** | Drowsiness Detection, Vision Transformers, Mood Estimation, Mobile Application, Artificial Intelligence(AI)/Machine Learning(ML) |
| **Data Availability** | Public<br>Yawning and Distracted Driver Dataset. YawDD (Yawning Detection Dataset) [6], Driver-Drowsiness-Dataset-D3S [10], National Tsing Hua University(NTHU) driver detection dataset [11] and private data collected when a user uses the mobile application. |
| **Metadata (Type of Data)** | Image, video |
| **Model Training and Fine-Tuning** | Train a generic VisionTransformer model on Drowsiness Detection. Fine-tune the same models for user customization. |
| **Testbeds or Pilot Deployments** | The system is currently in the research and prototype validation stage. While a formal testbed or public pilot deployment does not yet exist. |

# 2 Use Case Description

## 2.1 Description

Autonomous driving represents the inevitable trajectory of the automotive industry. However, the widespread adoption of fully autonomous technologies is not an immediate objective, nor will it occur uniformly across different regions. Consequently, existing Advanced Driver-assistance Systems (ADAS), such as Driver Monitoring Systems (DMS) technologies, remain highly relevant as they can anticipate and mitigate potential accident risks [1].

Driver drowsiness is a major contributor to road accidents globally, posing serious risks to both drivers and the public. According to the European Commission, fatigue-related crashes account for up to 20% of all road traffic accidents in Europe. In the United States, the National Highway Traffic Safety Administration (NHTSA) estimates that drowsy driving causes approximately 100,000 police-reported crashes annually, resulting in over 1,500 deaths and 71,000 injuries. One widely cited real-world incident occurred in 2014, when a Walmart truck driver—allegedly awake for more than 24 hours—caused a fatal crash involving comedian Tracy Morgan, bringing national attention to the dangers of fatigued driving [2].

In response to this critical safety issue, Driver Monitoring Systems (DMS) powered by AI have emerged as vital tools to detect drowsiness and prevent accidents. However, current systems often fall short in reliably identifying drowsiness across diverse populations due to biased training data and overly generalized models.

Upon evaluating a DMS system utilizing state-of-the-art facial landmark recognition models, it was observed that these models exhibit limitations in accurately identifying facial landmarks across individuals of diverse racial backgrounds and varying facial characteristics, such as the presence of facial hair [2-3]. This disparity may result in false negative outcomes in the drowsiness detection alerts, thereby compromising its reliability and effectiveness. The underlying causes of these limitations may include insufficient racial diversity in the training datasets and the inherent simplicity of the models, which may prioritize generalized facial features over nuanced individual variations. Addressing these challenges is essential to ensuring the inclusivity, accuracy, and equitable performance of DMS technologies across diverse populations.

In line with the principles of inclusivity, fairness, and the ethical deployment of Artificial Intelligence, we suggest a use case with native alternative data models designed specifically to address the identified limitations. These models have a strong emphasis on ethical considerations and equitable performance across diverse populations, ensuring that individuals of all backgrounds are accurately represented. Importantly, this commitment to native fairness and inclusivity will be pursued without compromising the overall efficacy and reliability of the model.

To evaluate the fairness and unbiasedness of our model we plan to use first of all some automated tools. Upon our initial research we encountered several automated tools with the most promising one being Fairlearn [7]. Fairlearn can be integrated seamlessly in our training pipeline and give us metrics on the predictions of the model such as demographic parity, predictive parity and calibration curves just to name a few key metrics to assess the fairness of a model [8]. We also plan to deploy automatic pipelines to check the bias of a dataset by identifying unbalanced in any dataset between the prediction classes and also creating a

race distribution if possible to determine that each race presented in the dataset is equally represented.

**Use Case Status:** The use case is in progress and part of a research project.

## 2.2 Benefits of the use case

Sustainable Cities and Communities: Public transit lies at the heart of efficient urban mobility, and keeping drivers alert is critical for accident prevention and rider confidence. By monitoring and mitigating fatigue and stress in real time, our app makes buses, trams, and trains noticeably safer—encouraging more commuters to choose mass transit over private cars and easing road congestion. When passengers know operators are actively supervised for drowsiness, they feel more secure, which is especially valuable for children, older adults, and people with disabilities who depend on these services. In doing so, the solution not only reduces crash risk but also fosters an inclusive network where everyone has equitable access. Finally, by boosting trust and ridership, this technology helps cut carbon emissions and supports the long-term vision of resilient, low-emission cities.

## 2.3 Future Work

Our proposed future work involves utilizing the app for data collection, to collect frames of the driver alongside physiological data, such as heart rate obtained from wearables, while driving. By gathering real-world data, we aim to enhance the accuracy and reliability of drowsiness detection. Additionally, we plan to enhance this dataset by ingesting information from already existing ones, to develop a more robust drowsiness detection model.

To ensure the proposed drowsiness detection system is both effective and equitable, we will adopt a multi-dimensional evaluation framework that includes metrics for accuracy, fairness, and latency. Accuracy will be measured using standard classification metrics such as precision, recall, F1-score, and area under the Receiver Operating Characteristic( ROC) curve Area Under the ROC Curve (AUC) on both general and user-specific datasets. Fairness will be evaluated through the use of tools like Fairlearn, which provides detailed metrics such as demographic parity, equal opportunity difference, and calibration curves across sensitive attributes (e.g., race, gender, facial features). These tools will be integrated directly into our model training pipeline for ongoing bias monitoring and mitigation.

In addition latency, which is critical for real-time mobile application deployment, will be assessed by measuring inference time on representative mobile hardware. The goal is to ensure sub-second predictions (<500 ms) to allow timely intervention without requiring constant cloud connectivity. Together, these validation methods will provide a robust and holistic view of model performance in real-world conditions.

A key focus of our research is ensuring that the dataset is as diverse as possible. This diversity is essential to eliminating racial and facial characteristic biases that often affect AI-driven detection systems. We plan to train different models on a wide range of facial structures, skin tones, and physical attributes, to create an unbiased system that provides fair and accurate results for all users. This approach ensures that drivers of different ethnic backgrounds are equally protected, promoting inclusivity in road safety technology.

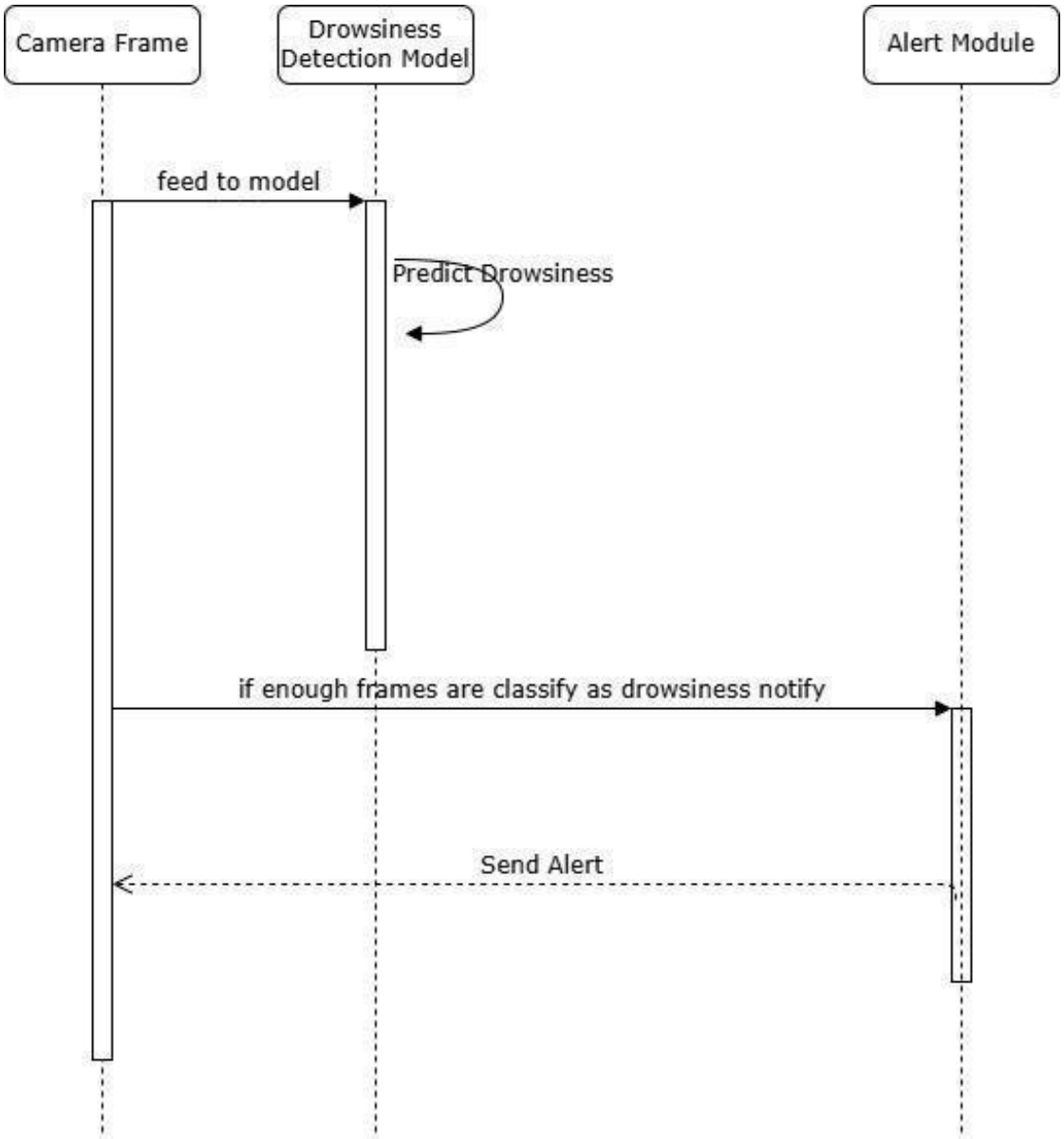
Another crucial aspect of our development is optimizing the drowsiness detection model for real-time performance, aiming to run it directly on the device, reducing reliance on cloud processing. This will ensure faster predictions and immediate alerts, enhancing driver safety without requiring constant internet connectivity. Real-time capabilities are critical in preventing accidents, as immediate intervention can prompt drivers to take necessary precautions before fatigue leads to dangerous situations.

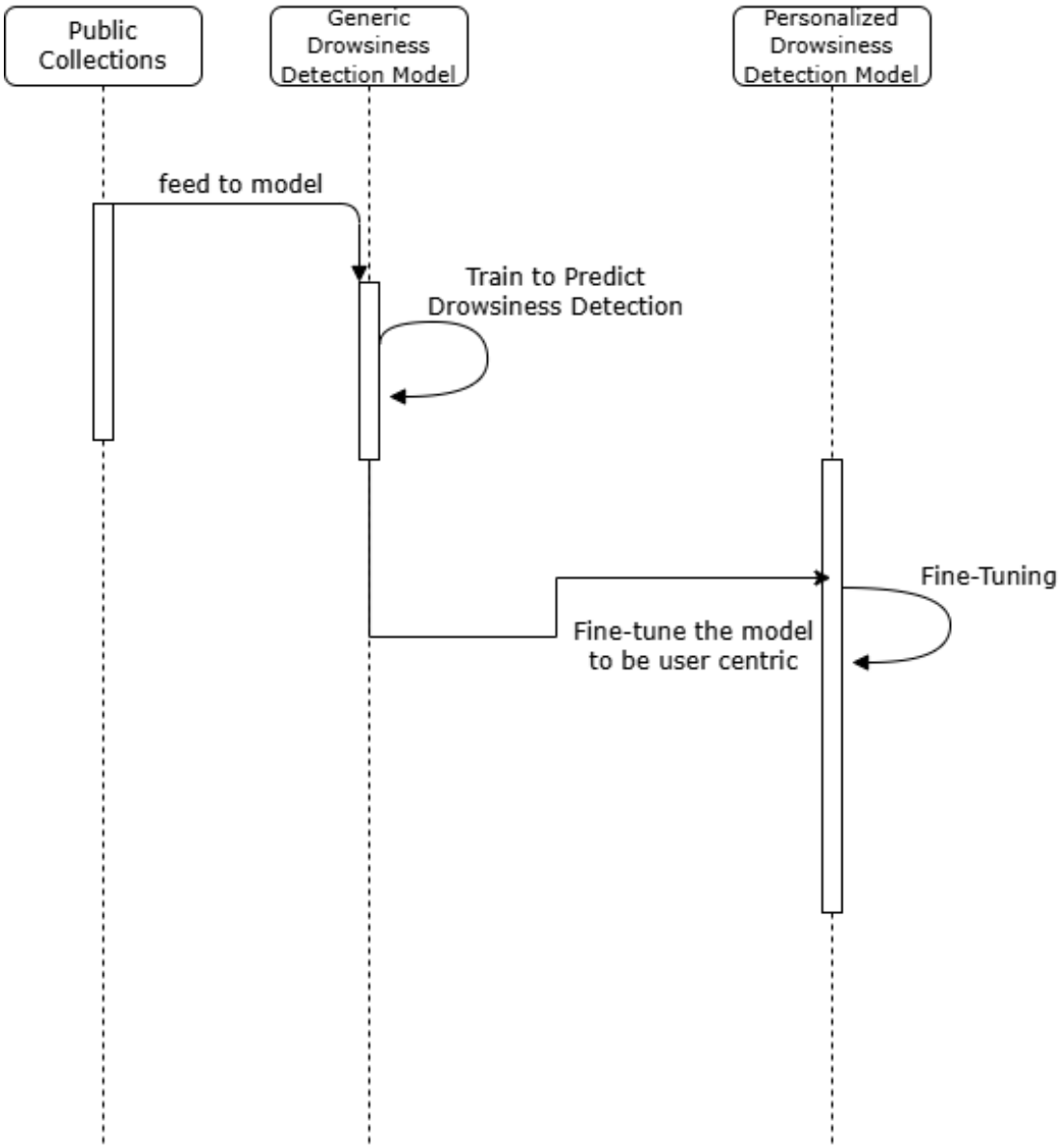
By prioritizing fairness, efficiency, and real-time responsiveness, our future work aims to create a cutting-edge AI-powered drowsiness detection system that is both inclusive and highly effective in ensuring road safety for all. By training multiple models we aim to provide the best model as a proof of concept that these biases can be mitigated in drowsiness detection.

## 3 Use Case Requirements

- **REQ-01**: Test the generated models offline on a diverse dataset containing different racial profiles and facial characteristics that will ensure that our model is unbiased and fair and doesn't lose performance.
- **REQ-02**: Train a generic model that can be fine-tuned and be customization. We plan on using VisionTransformer models as our basic architecture since they showed promising results on training on different generic tasks. This can ensure we have a base model and it can be fine-tuned to the user specific expression. That might make the model better on individual performance and it won't need a lot of data to be fine-tuned on the user expression.
- **REQ-03**: The model should be able to make predictions real-time on a mobile device. We want to use a mobile device which is widely available as our capture medium for the user's data. Our model should be able to make predictions in real-time when running on a mobile device.
- **REQ-04**: On top of our model being unbiased in terms of race and facial characteristics it will also need to be independent of factors such as the use of sunglasses. Also it should be independent of different weather and lighting settings.

## 4 Sequence Diagram






## 5 References

[1] International Transport Forum, OECD. (2020). The Future of Autonomous Driving: Challenges and Opportunities. [online] Available at: https://www.itf-oecd.org/

[2] https://www.bbc.com/news/entertainment-arts-33880232 Source of news about car accidents

[3] Liu, S., Li, H., and Choi, S. (2020). Deep Learning for Face Recognition: A Critical Analysis and Practical Guide. IEEE Transactions on Neural Networks and Learning Systems, 31(4), pp. 1313-1327.

[4] Buolamwini, J. and Gebru, T. (2018). Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification. Proceedings of the 1st Conference on Fairness, Accountability, and Transparency (FAT* '18)

[5] Gupta, I., Garg, N., Aggarwal, A., Nepalia, N., & Verma, B. (2018, August). Real-time driver's drowsiness monitoring based on dynamically varying threshold. In 2018 Eleventh International Conference on Contemporary Computing (IC3) (pp. 1-6). IEEE

[6] Processing;Convolutional Neural Networks;Faster R-CNN;Datasets;Raw Data},Shabnam Abtahi, Mona Omidyeganeh, Shervin Shirmohammadi, Behnoosh Hariri, August 1, 2020, "YawDD: Yawning Detection Dataset", IEEE Dataport, doi: https://dx.doi.org/10.21227/e1qm-hb90.

[7] Ghoddoosian, R., Galib, M. and Athitsos, V., 2019. A realistic dataset and baseline temporal model for early drowsiness detection. In Proceedings of the ieee/cvf conference on computer vision and pattern recognition workshops (pp. 0-0)

[8] Weerts, H., Dudík, M., Edgar, R., Jalali, A., Lutz, R. and Madaio, M., 2023. Fairlearn: Assessing and improving fairness of AI systems. Journal of Machine Learning Research, 24(257), pp.1-8.

[9] Pagano, T.P., Loureiro, R.B., Lisboa, F.V.N., Peixoto, R.M., Guimarães, G.A.S., Cruz, G.O.R., Araujo, M.M., Santos, L.L., Cruz, M.A.S., Oliveira, E.L.S., Winkler, I., Nascimento, E.G.S., 2023. Bias and Unfairness in Machine Learning Models: A Systematic Review on Datasets, Tools, Fairness Metrics, and Identification and Mitigation Methods. Big Data and Cognitive Computing 7, 15. https://doi.org/10.3390/bdcc7010015

[10] Gupta, Isha & Garg, Novesh & Aggarwal, Apoorva & Nepalia, Nitin & Verma, Bindu. (2018). Real-Time Driver's Drowsiness Monitoring Based on Dynamically Varying Threshold. 1-6. 10.1109/IC3.2018.8530651.

[11] Weng, Ching-Hua et al. "Driver Drowsiness Detection via a Hierarchical Temporal Deep Belief Network." ACCV Workshops (2016).

# Use Case 3: From Silence to Speech: A Novel AI System for Speech Development for the Deaf Through Lip-Reading

**Organization:** Huazhong University of Science and Technology

**Country**: China

**Contact Person(s)**:

Ran Wang, rex_wang@hust.edu.cn

Yang Xiao, Yang_Xiao@hust.edu.cn

## 1 Use Case Summary Table

| Item | Details |
|---|---|
| Category | Accessibility |
| Problem Addressed | By leveraging Artificial Intelligence(AI) technologies to enable individuals with hearing impairments to regain vocal communication capabilities, this initiative aims to mitigate educational resource inequities stemming from the scarcity of human and equipment resources. |
| Key Aspects of Solution | Develop an AI-powered lip-reading recognition system that converts visual facial signals into textual language information through system processing. This system, integrated with standardized video materials, will serve as a self-learning guide for deaf individuals to study lip language. The platform will be disseminated through online networks to ensure enough educational resources. |
| Technology Keywords | Deep Learning, Vision-language model |
| Data Availability | Public Data source of Lip Reading: CAS-VSR-W1k [1] |
| Metadata (Type of Data) | Videos |
| Model Training and Fine-Tuning | Our vision-language model was trained and tested on public video datasets. |
| Testbeds or Pilot Deployments | The AI algorithms are deployed on cloud-based servers, and the product is delivered to deaf users through client-side applications.<br>Link: [2] |

# 2 Use Case Description

## 2.1 Description

As a developing country, China has 27.8 million hearing-impaired individuals who struggle with bidirectional communication through sign language [31], making them a vulnerable group in society. Most deaf individuals physiologically retain speech capacity but gradually lose verbal abilities due to auditory deprivation- unable to hear, unwilling to speak, and ultimately incapable of articulation [42]. Therefore, we target hearing-impaired individuals with residual speech capacity, applied in daily social, educational, and rehabilitation scenarios to aid their vocal learning and social integration.

Lip-reading forms the foundation for vocal rehabilitation. However, the lack of commercial lip-reading tools forces deaf learners to rely solely on special education teachers in limited institutions. Severe shortages and uneven distribution of educational resources create significant barriers. For instance, Wuhan (capital of Hubei Province) hosts 15,334 hearing-impaired residents [53], yet its provincial rehabilitation center employs merely a dozen special educators.

This pioneering project leverages AI technology to empower deaf individuals in lip-reading acquisition and active vocalization, enabling bidirectional communication. We have developed the world's first AI-powered lip-reading training system, allowing autonomous learning through real-time visual feedback. Users imitate lip movements from instructional videos while our computer vision-based algorithm analyzes and evaluates their articulatory gestures, providing immediate corrective feedback. This system particularly benefits remote communities through web access, reducing dependence on human instructors according to field tests. To date, 343 deaf users have regained speech capabilities, with data showing 3× learning efficiency improvement. Our lip-reading recognition algorithm demonstrates exceptional performance in complex scenarios, achieving 10% higher accuracy than state-of-the-art methods on international benchmarks [64].

Currently deployed in rehabilitation centers and special schools across China through non-profit partnerships, the system remains free for users. Looking forward, we aim to harness AI's potential in ensuring equitable educational opportunities for all hearing-impaired individuals, empowering them to express, learn, and thrive equally in our digital society - transforming impossibility into reality. The system protects user rights and data privacy by anonymizing user information and implementing access control. Additionally, video data is only used for model inference on the server side and not stored long-term, ensuring compliance with data security standards. We are also promoting the localized deployment of models to achieve local processing priority, minimizing the need for transmitting sensitive lip-reading video data.

**Use Case Status:** The case is part of larger research projectArtificial Intelligence, Assist Disabled Communities

**Partners**: Huazhong University of Science and TechnologyN/A

## 2.2 Benefits of the use case

Quality Education: The project leverages AI technology to overcome the traditional dependence on special education teachers for lip-reading instruction. It enables deaf individuals—especially

those in remote or underserved areas—to learn lip-reading independently via online platforms. This helps address the critical shortage and uneven distribution of educational resources, contributing to inclusive and equitable access to quality education and lifelong learning opportunities.

By replacing conventional, human-led teaching methods with AI-driven tools, the solution significantly reduces educational costs while offering standardized, repeatable learning experiences. It also supports the advancement of education infrastructure and digital learning capabilities, particularly for learners in areas where special education resources have long been scarce.

Reduced Inequalities: The project tackles communication-related exclusion by equipping deaf individuals with AI-based tools to enhance their speech and lip-reading abilities. This supports more active participation in social life and helps reduce disparities caused by physiological challenges. With 343 individuals already enabled to speak using the system, it has proven effective in increasing confidence and enabling integration into broader society.

By offering free access to deaf schools and rehabilitation centers—especially in resource-limited regions—the initiative ensures that vulnerable populations receive targeted support. For example, deaf individuals in rural areas of China can independently engage in learning without depending on in-person instruction, narrowing the gap in access between urban and rural areas.

In contrast to traditional systems that focus solely on helping hearing people understand the deaf (e.g., sign language interpreters), this solution empowers deaf individuals to express themselves vocally, fostering two-way communication. This shift toward vocal empowerment plays a vital role in building a more inclusive and communication-accessible society.

## 2.3 Future Work

Data Collection: n:

We will expand Chinese lip-reading datasets to enhance AI model accuracy in Chinese visual speech recognition. This will improve environmental adaptability (e.g., strong backlighting/low-light conditions), reduce physical space requirements for lip-reading training, and ensure equal access to educational resources for deaf communities in underdeveloped regions.

Model Development: t:

Through lightweight model optimization for local deployment on mobile devices, we aim to:

- Reduce on-device feedback latency
- Minimize reliance on stable internet connections

This technological advancement will optimize usability in remote areas and low-signal scenarios, guaranteeing equitable access to efficient learning support.

Standardized Educational Framework: k:

We will refine system functionalities and pedagogical methods through practical implementation, ultimately establishing a standardized deaf education framework based on lip-reading and speech production. Through nationwide deployment via open-source system access and

optimized teaching protocols, this initiative will empower the deaf community to achieve lifelong learning and personal development.

Through continuous technological iterations and public benefit initiatives, we are dedicated to building a more inclusive and equitable global education ecosystem for the hearing-impaired community, fully realizing the SDG principle of "leaving no one behind."

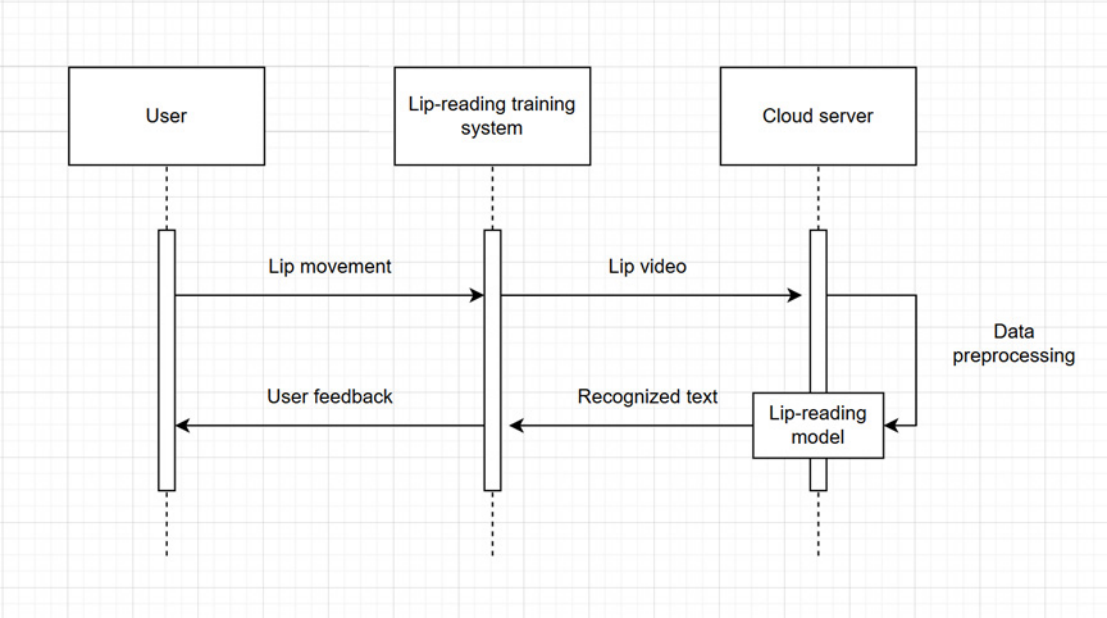
System Maintenance:

To enhance system reliability and user experience, we improve functional stability by addressing user feedback, refine lip-reading algorithms through analyzing real-world bad cases to boost accuracy, and conduct regular inspections of model deployments and server operations for sustained performance.

## 3 Use Case Requirements

- **REQ-001**: Users must possess network-enabled terminal devices equipped with video capture capabilities.
- **REQ-002**: The cloud-based deployment of the system requires sufficient computational power to support concurrent user access.
- **REQ-003**: The core functionality of lip-reading recognition lies in converting user captured video footage into textual recognition outputs.
- **REQ-004**: The AI system must efficiently and effectively organize, utilize, and expand training and testing datasets.

## 4 Sequence Diagram



## 5 References

[1] https://vipl.ict.ac.cn/resources/databases/201810/t20181017_32714.html: Data source of Lip Reading: CAS-VSR-W1k

[2] https://github.com/HUST-DPKW/Micro-lip-reading/blob/main/index.md: Brief introduction towards micro lip reading in the wild

[3] McKee, D., & McKee, R. L. (1992). What's so hard about learning ASL: Students & teachers' perceptions. Sign Language Studies, 75(1), 129-157. https://doi.org/10.1353/sls.1992.0000

[4] Geers, A. E., Mitchell, C. M., Warner- Czyz, A., Wang, N.-Y., & Eisenberg, L. S. (2016). Early sign language exposure and cochlear implantation benefits. Pediatrics, 138(1), e20163489. https://doi.org/10.1542/peds.2016-3489

[5] https://www.rmzxw.com.cn/c/2021-03-16/2810330.shtml: News about hearing screening for hearing impaired people in Wuhan, Hubei

[6] https://ieeexplore.ieee.org/document/10876180: Toward Micro Lip Reading in the Wild: Dataset, Theory, and Practices

# Use Case 4: On-Device Intelligence Empowers People with Disabilities and Smart Interaction Scenarios

**Organization**: vivo

**Country**: China

**Contact Person**:

Mengzhu Li, limengzhu.ai@vivo.com

## 1 Use Case Summary Table

| Item | Details |
|---|---|
| Category | Accessibility |
| Problem Addressed | China has nearly 50 million individuals with hearing or visual impairments. This case aims to leverage the on-device AI capabilities to provide intelligent assistive tools for people with disabilities and multilingual cultural groups, optimizing their experiences in daily life and specific smart scenarios. |
| Key Aspects of Solution | On-device AI capabilities create inclusive assistive technologies targeting accessibility challenges. By deploying localized lightweight AI models (e.g., vivo Voice, vivo Sight), the project enables real-time, privacy-centric tools for hearing/visual impairments and multilingual communication, eliminating cloud dependency. Innovations include multimodal interaction and offline-first functionality to ensure reliability without network constraints.<br>1. On-Device Lightweight AI Models & Privacy-by-Design Architecture<br>2. Full-Spectrum Accessibility Features:<br>• Real-time conversion between voice and text for smoother communication.<br>• Real-time environmental Q&A in life.<br>• Sign Language Translator: The first application in China to implement sign language recognition technology [1]. Sign language experts are engaged throughout the product lifecycle.<br>• Score Reading: Chase your music dream and savor the joy.)<br>3. Offline-First Design:<br>Critical functions (e.g., speech/text conversion, image recognition) operate without internet connectivity, serving remote areas and emergencies.<br>4. Ecosystem Collaboration, Launched Open Developer Platform to foster third-party innovations. The platform provides different Application Programming Interfaces(APIs) to developers such as real-time recognition, general Optical Character Recognition(OCR), Text-to-Speech(TTS) and vivo-Blue Language Model(LM) [2], etc. |
| Technology Keywords | Automatic Speech Recognition, Facial Recognition, Multi-target Tracking/Recognition [3], Multi-modal Large Model, Music Note Recognition Technology, Audio Event Detection, Sign Language Recognition Technology,Sign Language Synthesis,Virtual Avatar, Offline Audio Recognition |

(continued)

| Item | Details |
|---|---|
| Data Availability | 1. National and international public data<br>• laion5B:[4]<br>• Taisu:[5]<br>• wukong:[6]<br>2. Internal company data<br>3. Third-party procurement data<br>4. User-authorized data<br>5. Generated data |
| Metadata (Type of Data) | Audio, Video and Image |
| Model Training and Fine-Tuning | Image caption provides a text description of an image. Visual Question Answering combines images and questions to predict answers.<br>Audio-visual speech recognition combines sound and video information to identify speech content.<br>The project focuses on model fine-tuning based on different scenarios. |
| Testbeds or Pilot Deployments | Project Accessibility Features: [7]<br>vivo origin OS: [8] |

# 2 Use Case Description

## 2.1 Description

### Context and Background

People with disabilities and multilingual cultural groups face significant challenges when using smart devices due to:

- Limited Functionality: Existing assistive technologies (e.g., speech-to-text, image recognition) heavily rely on cloud-based AI processing, leading to high latency and offline inaccessibility.
- Privacy Risks: Cloud data transmission increases vulnerabilities to personal information leaks, especially for sensitive data of disabled users.
- High Energy Consumption: Cloud computing's resource-intensive nature conflicts with sustainable development goals.

These technological barriers exacerbate difficulties in communication, information access, and environmental adaptation, directly impacting quality of life.

## Objectives and Aims

This project leverages on-device AI technology to:

- Develop real-time, secure, and efficient assistive tools addressing core needs of individuals with hearing/visual impairments and multilingual groups (e.g., voice interaction, sign language translation, offline support).
- Break Technical Barriers: Reduce cloud dependency through localized AI processing, ensuring privacy and energy efficiency, and advancing the "AI for Good" vision.
- Promote Social Equity: Bridge the digital divide for disabled communities, enhancing social participation and life autonomy.

Core Technological Breakthroughs:

- On-Device Lightweight AI Models: Integrate deep learning with edge computing to optimize efficiency across hardware configurations (e.g., vivo Voice, vivo Sight).
- Multimodal Fusion: Combine voice, image, and tactile perception (e.g., sign language recognition, screen reading) for comprehensive accessibility.
- Offline-First Design: Enable critical functions (e.g., Sound Recognition, Offline Audio Recognition) without network connectivity.

Key Features:

- Voice/Accessibility Calls: Real-time speech-to-text and multilingual translation for seamless communication.
- Sign Language Translator: China's first AI-powered tool converting text to continuous sign language via virtual avatars. Sign language experts are engaged throughout the product lifecycle. During data collection, they assess data accuracy and practicality, and also help validate standards, protocols, and data usability. In test/launch phase, they provide recommendations and product experience reports. vivo Sight: Visual multimodal technology providing environmental descriptions and image Q&A for visually impaired users.

Social Impact:

- Reduced Inequality: Serve over 500 million global users (300 million in China), empowering 80 million individuals with disabilities and 300 million elderly users through digital literacy and employment skills.
- Sustainable Production: On-device AI reduces cloud energy consumption by 90%, advancing green computing and privacy protection.

Industry Influence:

- Standardization: Collaborate with governments and Non-Governmental Organizations(NGOs) to establish accessibility technology standards.
- Ecosystem Collaboration: Launch an open developer platform to foster third-party innovations.
- Technological Evolution: Continuously upgrade multimodal models for real-time interaction and predictive accuracy, building an inclusive "Information Accessibility" ecosystem.

Impact on healthcare:

Reduced Inequality

By providing AI-assisted features on the device, the use case provides more equal opportunities for access to technological services for disable groups (such as people with visual or hearing impairments) and multilingual communities. On-device AI capabilities, such as real-time speech-to-text, image recognition description, and sign language recognition, help them interact with the world more conveniently, participate in social activities, and access information and resources. Due to the widespread use of smartphones, the features can benefit a broader user group, narrowing the digital divide, fostering social inclusivity, and reducing inequalities caused by disabilities or impairments.

Responsible Consumption and Production

The on-device AI approach reduces dependence on cloud servers, minimizing the energy consumption required for data transmission and processing, aligning with the green and environmentally friendly principles. This low-energy, high-efficiency computing method reduces carbon emissions and supports sustainable development. Moreover, protecting user privacy and preventing data abuse reflects the commitment to consumer rights. By providing users with sustainable products and services, we advocate for more environmentally friendly and responsible consumption model.

**Use Case Status:** The use case is part of a larger product development

**Partners:** China Foundation for Disabled Persons(CFDP) [9], China Association of Persons with Visual Disabilities (CAPVD), Audiology Development Foundation of China(ADFC)

## 2.2 Benefits of the use case

Reduced Inequality: By offering AI-assisted features directly on mobile devices, the use case enhances access to technology for underserved and marginalized groups, including individuals with visual or hearing impairments and speakers of diverse languages. On-device AI functionalities—such as real-time speech-to-text conversion, image recognition with descriptive output, and sign language recognition—empower users to engage more easily with their environment, participate in social life, and access critical information. Leveraging the widespread availability of smartphones, these features reach a broader population, help bridge the digital divide, and promote greater social inclusion by addressing barriers caused by disability or language limitations.

Responsible Consumption and Production: The use of on-device AI reduces reliance on cloud-based computing, thereby lowering the energy consumption associated with data transmission and remote processing. This approach promotes environmental sustainability through energy-efficient computing, ultimately contributing to lower carbon emissions. Additionally, by enabling local data processing, the system upholds user privacy and helps prevent misuse of personal data. These measures support a more responsible and sustainable technology consumption model that aligns with environmentally friendly and user-respecting design principles.

## 2.3 Future Work

Enhance the application effects of on-device AI in empowering disabled individuals and smart scenario and accelerate the process of sustainable development internationalization.

1. Technology Optimization and Product Innovation

Optimize on-device AI algorithms, focusing on improving model lightweighting and efficiency, adapting to mobile phones with different hardware configurations. Integrate the latest technologies combining deep learning and edge computing to enhance AI real-time responses and prediction accuracy. Explore multimodal AI technologies, integrating voice, image, and tactile perception to provide more comprehensive assistive features. On this basis, develop more accessible intelligent features to meet the needs of various disabilities and scenarios, enhancing the user experience.

2. Collaboration and Ecosystem Building

Strengthen cooperation with governments, research institutions, NGOs, and other enterprises 10to jointly promote the development and application of accessibility technologies. Participate in the development of industry standards to facilitate the standardization and popularization of AI in the accessibility field. Additionally, build an open developer platform to encourage innovative applications based on AI platform.

3. Internationalization and Market Expansion

Optimize AI solutions based on the accessibility needs and regulatory requirements of different regions. Actively participate in international exhibitions and events related to AI and accessibility to enhance the benefit of AI in the use case.

4. Security and Privacy Protection

Enhance on-device AI security and privacy protections to safeguard user data locally [11]. Increase the investment in AI security research and privacy protection algorithms to build user trust.

# 3 Use Case Requirements

- **REQ-01**: Algorithm Capabilities

  The multimodal large model uses Large Language Model(LLM) as its base and adds a visual module to LLM, enabling the model to simultaneously process data from both text and image modalities, which can better handle complex and real-world problems.

- **REQ-02**: Hardware Requirements

  Users are required to use smartphones that support accessibility features to ensure that the devices can run high-performance algorithms and handle complex tasks.

- **REQ-03**: User Feedback Collection and Analysis Technology

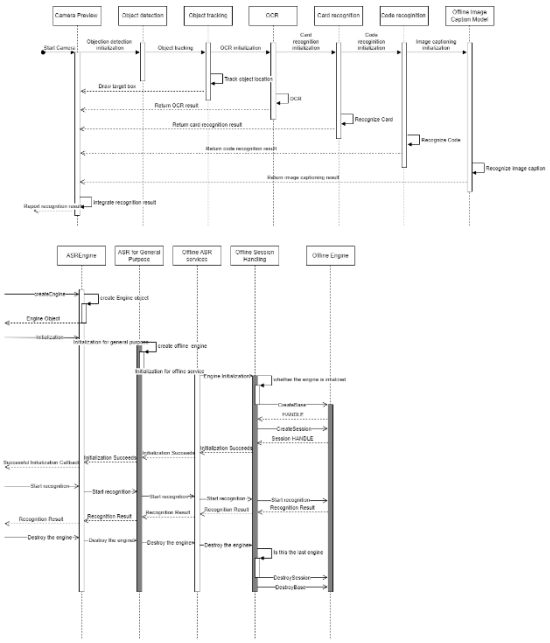
  By collecting user feedback on the accessibility feature via online and offline channels, we use data analysis technology to evaluate these feedbacks and optimize the accessibility feature accordingly. Such an analytical method aids in understanding user needs and behaviors, thus enhancing feature design and boosting user satisfaction.

- **REQ-04**: Social Development Trends and Corporate Responsibility

Given the aging population and the growing demand for accessibility in social, accessible products are necessary. This enables AI to permeate all aspects of life through mobile phones, supporting the quality of life for users with visual and hearing impairments.

In order to meet the above use case requirements, this accessibility project combines technology, hardware, requirements, and corporate responsibility to provide high-quality accessibility services for visually and hearing-impaired users, promoting inclusive development in society.

## 4 Sequence Diagram




## 5 References

[1] J. He et al., "vivo Sign Language Avatar Translation System," vivoMobileCommunicationCo.,Ltd, 2024. Available: https://aclanthology.org/2024.ccl-3.43.pdf

[2] X. Lu et al., "BlueLM-V-3B: Algorithm and System Co-Design for Multimodal Large Language Models on Mobile Devices." arXiv, Nov. 16, 2024. doi: 10.48550/arXiv.2411.10640. Available: http://arxiv.org/abs/2411.10640.

[3] L. Xu et al., "A Novel Feature Based on Graph Signal Processing for Detection of Physical Access Attacks," presented at the Proc. Odyssey 2022, 2022, pp. 107–111. doi: 10.21437/Odyssey.2022-15. Available: https://www.isca-archive.org/odyssey_2022/xu22_odyssey.html#:~:text=With%20this%20motivation%2C%20we%20propose%20a%20novel%20feature,the%20graph%20Fourier%20analysis%20to%20extract%20GFCC%20features.

[4] Multimodal Large Model Training Database:laion5B: https://laion.ai/blog/laion-5b/

[5] Multimodal Large Model Training Database:Taisu: https://github.com/ksOAn6g5/TaiSu

[6] Multimodal Large Model Training Database: wukong:https://wukong-dataset.github.io/wukong-dataset/index.html

[7] https://accessibility.vivo.com/?from=CSR_HomePage

[8] https://www.vivo.com.cn/originos: vivo origin OS

[9] vivo and China Disabled Persons' Federation reach strategic cooperation to promote technology to help people with disabilities and share a better life.May 16,2024.Available: https://www.vivo.com.cn/brand/news/detail?id=1245[Accessed: May 28, 2025]

[10] vivo On-Device AI Accessibility Development Guide 2024

[11] vivo Privacy Center: https://privacy.vivo.com/overview

# Use Case 5: AI-Powered Behavior Recognition for Developmental Disability

**Organization**: SK telecom

**Country:** South Korea

**Contact Person:**

Choong Hwan Choi, ch.choi@sk.com

## 1 Use Case Summary Table

| Item | Details |
|---|---|
| Category | Accessibility |
| Problem Addressed | Manual behavior tracking in care centers is labor-intensive, inconsistent, and delays intervention for individuals with developmental disabilities. Caregivers are overburdened and lack data to support precise behavior intervention planning. |
| Key Aspects of Solution | Detects challenging behaviors in real-time using Vision Artificial Intelligence (AI). The solution runs on edge devices with lightweight model and personalizes recognition using face and clothing. Finally it provides statistics and reports for intervention planning. |
| Technology Keywords | Vision AI, Behavior Recognition, Edge AI, Knowledge Distillation, Pose Estimation,Re-identification (Re-ID), Metadata Analytics |
| Data Availability | Private (collected from care centers, actor-based simulations, and generated by Generative AI (gen-AI); labeled by professionals and social enterprise). |
| Metadata (Type of Data) | Video frames, pose keypoints, time/location metadata, user annotations, behavior labels. |
| Model Training and Fine-Tuning | Ensemble model using image, keypoint, and video streams.<br>Trained with 10M+ frames.<br>Techniques: attention, multi-stream fusion, knowledge distillation (using SK Telecom's VL2Lite framework).[1]. |
| Testbeds or Pilot Deployments | Deployed in 10+ care centers in South Korea (Seoul, etc.) [2]. |

# 2 Use Case Description

## 2.1 Description

CareVia is an AI-powered behavior recognition solution designed to support individuals with developmental disabilities. In South Korea, many people with autism spectrum disorders face significant safety risks due to a lack of sufficient care infrastructure. The leading cause of death among individuals with severe autism is accidents, and most care centers are critically understaffed.

CareVia uses SK Telecom's Vision AI to detect and analyze challenging behaviors, such as self-harm, aggression, and wandering. The system automates behavior tracking that was traditionally recorded manually by caregivers, helping to reduce workload and improve intervention planning by professionals. The AI runs in real-time on edge devices, utilizing lightweight models to ensure fast and accurate responses in environments like care centers, hospitals, and special schools.

This solution is the first of its kind in Korea, tailored specifically to behavioral monitoring for individuals with developmental disabilities. By automating data collection and offering visual analytics, it empowers care professionals and families to deliver timely and personalized support. CareVia has been deployed in 12 centers across South Korea, with the following observed outcomes:

### Expected Impact

1. Enhances safety and well-being of individuals with developmental disabilities.
2. Reduces administrative burden on caregivers.
3. Improves accuracy and speed of intervention planning.
4. Behavior tracking workload reduced from ~3 hours to 5 minutes per session.
5. AI accuracy of 85%+ in real-time behavior recognition.
6. Self-harm incidents decreased by 76.5% [3] after intervention (clinical study with Severance Hospital, 2024).
7. Positive feedback from caregivers who report improved documentation, reduced stress, and better intervention planning.

These results were also presented at the American Academy of Pediatrics Conference (Oct 2024) and supported by real-world behavior intervention specialists.

**Use Case Status:** CareVia uses SK Telecom's Vision AI to detect and analyze challenging behaviors, such as self-harm, aggression, and wandering.

**Partners: SK Telecom, Seoul National University Hospital, Focus AI, Korea ABA Behavioral Development Institute, AIWORKX, local governments (e.g., Seoul).**

## 2.2 Benefits of the use case

Good Health and Well-being: The solution enhances access to healthcare and improves the quality of services for vulnerable groups, particularly individuals with developmental disabilities. By enabling timely interventions, it supports better health outcomes and overall well-being.

Industry, Innovation, and Infrastructure: The use case leverages advanced AI and scalable systems to modernize care delivery. By utilizing existing CCTV infrastructure, it introduces innovative solutions without the need for costly hardware, making implementation both practical and efficient.

Reduced Inequalities: By offering low-cost, AI-powered care tools in under-resourced care facilities, the solution helps bridge service gaps and promotes equity. It ensures that individuals in underserved areas benefit from the same quality of care and digital support as those in more developed settings, reducing disparities and fostering inclusion [4].

## 2.3 Future Work

- Model enhancement: Integrate audio and biosensor (smart watch/band, Internet of Things (IoT) devices) data for multimodal behavior analysis.
- New applications: Expand to elderly care, childcare, and school safety.
- Continual learning: Improve accuracy by retraining with new data.
- Deployment at scale: Extend rollout through national and global public-private partnerships.
- Collaborations: Align with International Telecommunication Union (ITU) efforts on AI standards for healthcare and safety monitoring.[5]
- Sharing data: Currently, due to the sensitive nature of the data (involving individuals with developmental disabilities), our datasets are private and used under strict ethical supervision.

# 3 Use Case Requirements

**REQ-01**: It is critical that the system deploys a lightweight AI model optimized for edge computing, capable of operating on devices with a maximum computational capacity of 100 TOPS (Tera Operations Per Second) or lower.

**REQ-02**: It is critical that the system performs real-time video analysis with behavior detection accuracy > 85% to ensure reliable monitoring and intervention capabilities..

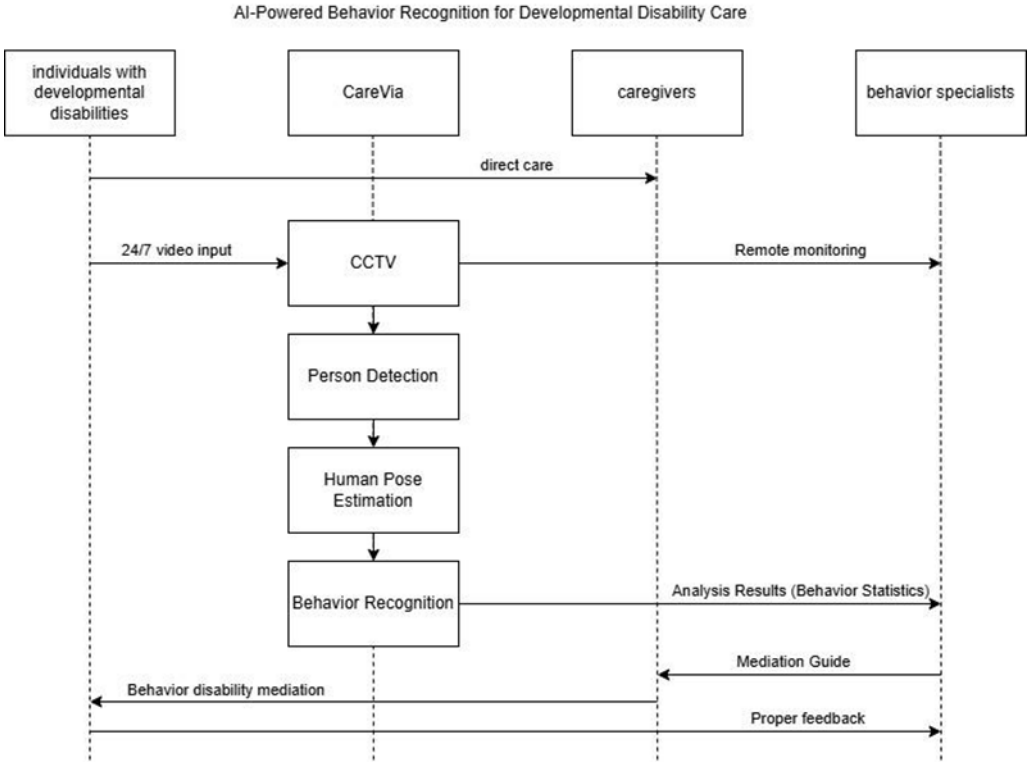
**REQ-03**: Support for privacy-preserving behavior monitoring (meta info only).

- CareVia is designed with privacy by design principles. The system does not store or transmit raw video externally.
- Anonymization is applied at the edge level via facial blurring and meta-information extraction.
- All processing stays on-device (within the AI Box) unless users explicitly export anonymized reports.
- The platform complies with local privacy laws and follows European Union (EU)-inspired data minimization principles.
- The system only shares timestamped behavior types and non-identifiable metadata across the platform.
- This ensures privacy is protected, even in environments where closed-circuit television (CCTV) use is highly regulated.

**REQ-04**: It is expected that the system seamlessly integrates with existing CCTV infrastructure to minimize deployment costs and leverage legacy hardware.

**REQ-05**: It is of added value that the system provides a web-based dashboard for behavior review, annotation, and intervention planning, enabling healthcare professionals to efficiently analyze and act on behavioral insights.

## 4 Sequence Diagram




## 5 References

[1] SK Telecom. (2025). VL2Lite: Task-specific knowledge distillation from large vision-language models to lightweight networks. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR).

[2] Pilot Deployments: http://seculab.asuscomm.com:30140/ Testbed portal

[3] World Health Organization, & World Bank. (2011). World report on disability. World Health Organization. https://www.who.int/publications/i/item/9789241564182

[4] United Nations. (n.d.). Sustainable development goals. https://sdgs.un.org/goals

International Telecommunication Union. (2019). Y.3172: Architectural framework for machine learning in future networks including IMT-2020. ITU-T. https://www.itu.int/rec/T-REC-Y.3172-201906-I

[5] International Telecommunication Union. (2022). F.747.17: Requirements and functional framework of AI-based behavior analysis in smart environments. ITU-T. https://www.itu.int/rec/T-REC-F.747.17-202203-I

# Use Case 6: Fine-Tuning ASR Language Models for Underrepresented Low-Resource Languages

**Organization**: United Nations High Commissioner for Refugees**(**UNHCR)

**Country**: Iraq

**Contact Person(s)**:

Roshna Abdulrahman, abdulrar@unhcr.org

Sofia Kyriazi, kyriazis@unhcr.org

Rebeca Moreno Jimenez, morenoji@unhcr.org

## 1 Use Case Summary Table

| Item | Details |
|---|---|
| Category | Accessibility |
| Problem Addressed | In the Kurdistan Region, UNHCR receives a lot of information from communities via unstructured text and audio recordings in Kurdish. These have to be manually processed and assessed, which is time-consuming, so the data is often not fully analyzed. |
| Key Aspects of Solution | The result of this project will be a language model for Kurdish speech to text - STT that is integrated into Kobotoolbox, a survey tool used by humanitarian organizations globally. Using this tool, colleagues can have more time to work on analyzing the data instead of spending time on transcription. |
| Technology Keyword | LLMs, GenAI, GPTs, Data Collection, Data Analysis, Low resource languages, ASR, STT, NLP |
| Data Availability | Public: a multi-language, open-source voice dataset for training speech-enabled applications [2]<br>An experimental dataset of Sorani Kurdish that could be used in speech recognition using CMUSphinx [1].<br>Private data collection:<br>• Synthetic Storytelling recordings<br>• Interviews with Kurdish Speakers |

(continued)

| Item | Details |
|---|---|
| Meta Data | Audio, Text (raw, processed)<br>This project develops a language model to turn audio of Kurdish speakers into text. The approach involves several sub-activities, including domain specific data collection, data validation and review, pre-processing, alignment of audio and transcript, reviewing selected technology, language model development and fine-tuning, and language model testing and evaluation.<br>Recent speech-to-text systems often use advanced AI/ML technologies such as Transformer models and Natural Language Processing (NLP) techniques. Transformer models, like those used in OpenAI's Whisper and Meta's Wav2Vec2.0, are popular due to their ability to handle long-range dependencies in data. NLP techniques are crucial for understanding and processing the linguistic aspects of speech. |
| Model Training and Fine-tuning | To address the challenges of low-resource languages, the project leverages multilingual models pre-trained on a wide range of languages, which can improve performance on low-resource languages. Data augmentation techniques, such as adding noise, changing pitch, or speed, are used to create more training data.<br>Including publicly available data, such as Mozilla's Common Voice, can provide a diverse set of voice samples that grow the dataset for training and fine-tuning the language model. Fine-tuning involves using frameworks like Hugging Face's Transformers to adapt pre-trained models to the specific dataset. We will ensure data privacy is maintained by running training locally.<br>By integrating these technologies and approaches, the project aims to develop an efficient and effective speech-to-text model for Kurdish and its dialects, enhancing data processing efficiency and contributing to industry innovation and infrastructure development |
| Testbeds of Pilot Deployment | • This model is a fine-tuned version of openai/whisper-small. [3]<br>• This model is a fine-tuned version of facebook/w2v-bert-2.0 on the common_voice_16_0 dataset [4]. |
| Code Repositories | See above. |

# 2 Use Case Description

## 2.1 Description

In the Kurdistan Region of Iraq and neighboring countries such as Syria, Turkiye, and Iran, as well as among Kurdish immigrants in Europe, various dialects of the Kurdish language are spoken. The UNHCR collects information from these communities through unstructured text and interview audio recordings in Kurdish. The manual processing and assessment of this data are time-consuming and often result in incomplete analysis.

The primary aim of this project is to develop a Kurdish speech-to-text language model integrated into KoboToolbox. This will enable UNHCR colleagues to focus on data analysis rather than

transcription. Future applications include summarizing, categorizing, or translating responses using other natural language processing tools. The specific objectives are:

- Collect humanitarian domain-specific audio data with transcription.
- Develop a language model with speech-to-text capabilities that can detect dialects and produce high-quality transcriptions.
- Create a functioning prototype in KoboToolbox.
- Conduct user testing and human verifications for a validation set to prove the model's accuracy.
- Collaborate with local academia and include underrepresented refugee communities.
- Document the process, lessons learned, and findings to create a replicable approach for low-resource language models.
- Ensure data protection, data responsibility, human rights, and ethics through the inclusion of UNHCR Iraq experts.

The project aims to create a template for low-resource languages in domain-specific fields by fine-tuning existing speech-to-text models using publicly available and domain-specific data. It addresses a gap in processing unstructured audio data by developing a language model for Kurdish and its dialects (Sorani and Kurmanji). This approach enhances data processing efficiency and contributes to industry innovation and infrastructure development. By integrating the model into KoboToolbox, the project will facilitate the collection and transcription of audio data, making the process more efficient and effective.

The expected impact and benefits of this solution include:

- Reduce manual processing of audio data in humanitarian domain.
- Improve response times to critical issues in emergency situations.
- Serve as a template for addressing similar challenges in other low-resource languages.
- Enable cross-validation of audio data.
- Generate dashboards and access previously unprocessed insights.
- Improve decision-making for resource allocation.
- Ensure replicability for other languages and operations

**Use Case Status:** It's in progress.

**Partners:** Kobo organization, University of Kurdistan - Hewler

## 2.2 Benefits of the use case

### Industry, Innovation and Infrastructure

This project contributes to innovation by establishing a framework for enabling low-resource languages to function in specialized domains, such as humanitarian audio data. Through the development of a language model tailored for Kurdish and its major dialects (Sorani and Kurmanji), it addresses a significant gap in digital infrastructure for processing unstructured voice data. The project fine-tunes existing automatic speech recognition (ASR) models using both publicly available and domain-specific data, improving the efficiency of data processing and promoting technological advancement in underrepresented linguistic communities.

### Reduced Inequality

By creating tools that can process and interpret audio data in Kurdish, the project helps ensure that feedback and interviews from Kurdish-speaking populations are accurately captured and understood. Transcribing voice into text allows for easier translation, classification, summarization, and integration into data systems such as dashboards and reports. This amplifies the voices of marginalized communities and enables more inclusive decision-making processes, ultimately supporting equity and representation.

### Partnerships and Collaboration

The project is grounded in collaboration with key partners to maximize impact. One partnership is with a local university that specializes in Kurdish computational linguistics, providing deep academic insight and cultural relevance. Another is with Kobo, the organization behind KoboToolbox, a widely used platform for data collection and visualization. By working with these partners, the project ensures the practical deployment of the language model within humanitarian data workflows and promotes adoption at scale through trusted platforms.

## 2.3 Future Work

Direct next steps:

- Conduct Data Collection: We will commission 10 hours of synthetic data and conduct humanitarian domain-specific audio data collection using the KoboToolbox survey.
- Development of Kurdish Speech to Text Language Model: We will set up pipelines for data processing, develop and fine-tune the speech-to-text model, and evaluate the model with test data and real-life testing.
- KoboToolbox Integration: We will create an API and establish a server for hosting language models, and modify KoboToolbox software to include the custom language model.
- Test the Integrated Model: We will test the integrated model in real-life settings and capture feedback.
- Communication and Dissemination: We will prepare a summary report and host a dissemination conference.

Further future work:

- Expanding the model to support additional Kurdish dialects beyond Sorani and Kurmanji, ensuring broader coverage and usability across different regions.
- Exploring the development of a Kurdish Large Language Model (LLM) capable of performing basic tasks such as summarization, translation, and Q&A with the transcribed text. This would enhance the utility of the transcriptions and provide more comprehensive language support.
- Collaborating with other operations to adapt the developed template for other low-resource languages. This would involve customizing the model and processes to fit the specific linguistic and operational needs of different regions and languages.

Our project has the potential for several collaborations and expansions:

- Partnering with other international organizations working in similar fields to share knowledge, resources, and best practices. This can help in scaling the project to other regions and similar languages.

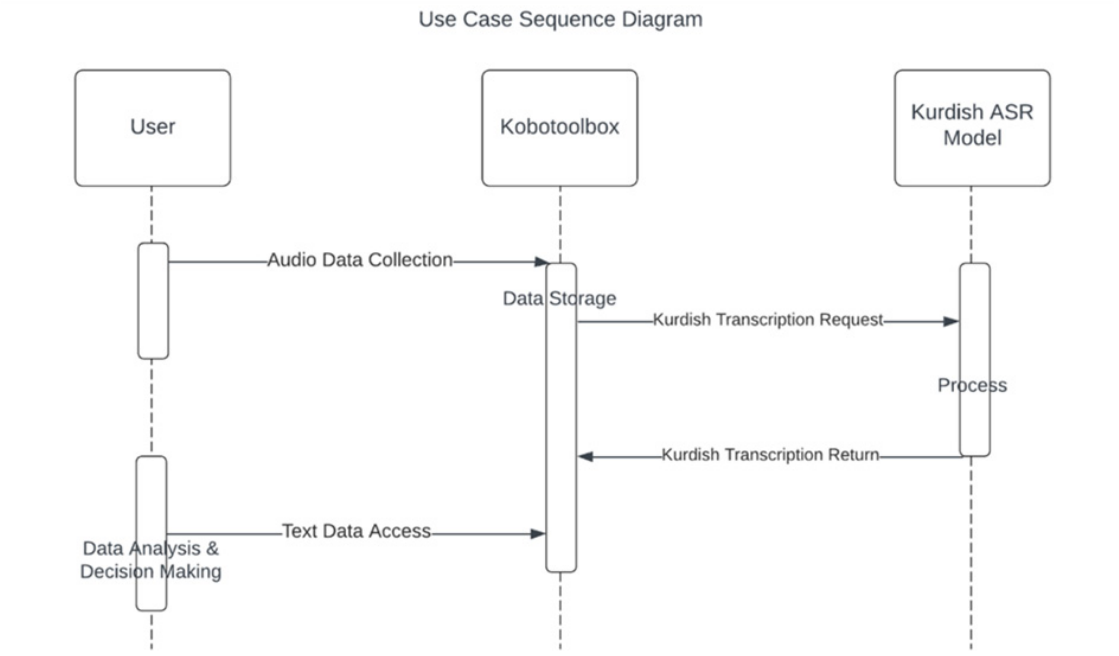
- Engaging with refugee communities and local stakeholders to ensure the project meets their needs and includes their feedback.

Working with technology providers like OpenAI or Meta to leverage their expertise and resources for model development and fine-tuning.

## 3 Use Case Requirements

- **REQ-01**: It is critical that sufficient processing power and compute time are allocated to support the fine-tuning of OpenAI's Whisper Large model. Given the training data exceeds 400 hours of audio, high-performance GPUs such as the RTX 4080 are required, with an estimated compute time of 200–400 hours depending on the number of training epochs.
- **REQ-02**: It is expected that dedicated technology infrastructure is established to host and serve the fine-tuned language models. This includes setting up a secure and scalable server environment with appropriate access configurations and safeguards to ensure reliable system availability.
- **REQ-03**: It is expected that model performance is continuously improved through the use of advanced machine learning techniques. This includes experimenting with larger transformer architectures and multilingual pre-trained models to enhance language understanding and accuracy.
- **REQ-04**: It is critical that the project is supported by a skilled interdisciplinary team. Required roles include NLP specialists, assistant researchers, software developers (front-end and back-end), User Interface/User Experience(UI/UX) designers, and software Quality Assurance(QA) testers to ensure robust development, integration, and user experience.

## 4 Sequence Diagram



## 5 References

[1] commonvoice.mozilla.org/en/datasets: This page serves as a reference hub for other open-source voice datasets and a home for our release updates as Common Voice grows.

[2] https://huggingface.co/roshna-omer/whisper-small-Kurdish-Sorani-10: This model is a fine-tuned version of openai/whisper-small on the None dataset.

[3] https://huggingface.co/roshna-omer/w2v-bert-2.0-krd-colab-CV16.0: This model is a fine-tuned version of facebook/w2v-bert-2.0 on the common_voice_16_0 dataset.

# Use Case 7: Blind Metaverse

**Organization**: Humanities Artificial Intelligence Research Center of Nara Knowledge Information Co.

**Country**: South Korea

**Contact**:

lee jae hyuk Principal researcher, jakelee@narainformation.com

shin dae geon Research Engineer, eorjs5518@narainformation.com

## 1 Use case summary table

| Item | Details |
|---|---|
| Category | Accessibility |
| The problem to be addressed | Lack of metaverse accessibility for the visually impaired, difficulty acquiring information, difficulty maneuvering, and complexity of User Interface(UI)/User Experience(UX). |
| Key aspects of the solution | Textualize information by touch and provide it out loud, detect surrounding information centered on your avatar, convert surrounding assets and event information into tactile (Braille) information, and reposition your avatar using tilt. |
| Technology keywords | One Touch UI, Gyro Sensor, Avatar Radar |
| Data availability | Private |
| Metadata (type of data) | Image |
| Model Training and fine-tuning | Blind-friendly UI/UX, using distance values to provide information, and tilt detection for avatar movement |
| Testbeds or pilot deployments | Avatar Radar-based Metaverse Accessibility Platform [1] |

## 2 Use Case Description

### 2.1 Description

The Metaverse opens up communication and learning opportunities for people with limited mobility beyond physical constraints. However, smartphones and tablets that are controlled by touching the display are difficult to access for the visually impaired, as the environment and UI are display-based. Avatar Radar, a touch UI for the visually impaired, and Avatar Radar, an avatar that moves according to the tilt of the display with a gyro sensor, are key features of Metaverse's accessibility technology, providing an optimized Metaverse experience for the

visually impaired by enabling them to recognize surrounding content without relying on sight. Users can tilt the display to move the avatar in the desired direction.

Avatar Radar uses distance values to detect assets within the detection distance and lists the detected assets to alert the user about the surrounding information. The amount of information retrieved varies by detection stage, and the list is ordered by distance for UI convenience. When the alarm goes off, the Asset Listboard UI pops up where the user can check his/her surroundings, and he/she can select and check what information is available in East. All of this information and images can also be viewed in Braille, thanks to a Bluetooth connection to a tactile display device. This gives the visually impaired the same access to information as sighted people and the freedom to choose what they want to see.

The Unity Collider component that this app is based on defines the shape of an object for handling physical collisions. Since colliders are invisible, they don't need to be the exact same shape as the object's mesh, so you can add a BoxCollider 2D to the Object component and set the collider size X,Y to create a box-like physical wall over the obstacle. By using colliders to create physical walls on exhibits and structures, we can warn the avatar when it approaches an obstacle ("There's a wall.", "There's an exhibit.") and prevent entry. This allows the user to know whether an obstacle is present or not, regardless of their vision.

The platform is not an assistive tool for the blind, but rather a metaverse where blind and sighted people exist and share the same virtual space together. The display on the platform functions as a "manipulation tool" for the visually impaired. Utilizing the tablet's tilt sensor to move the avatar turns the display into an interface that can be manipulated like a joystick. This allows the user to take the initiative to explore and experience the space.It also allows sighted and non-sighted people to not only experience the content in the same virtual space, but also to share their experience and empathy with the exhibit through a simple means of feedback: liking it. Once again, shaking the exhibit's description pop-up will trigger a thumbs up, and you'll hear an audible indication of how many thumbs up the exhibit has received. This allows people to share their feelings about the same content, even if they encounter it in different ways.

**Use case status: This use case reflects the feedback after the blind planner participated in the accessibility planning and the blind person used it directly at the disabled festival exhibition.**

Partner: N/A

## 2.2 Benefits of the use case

Metaverse accessibility technologies empower users to engage in learning and exploration without relying on sight. These tools help eliminate marginalization and discrimination faced by individuals with visual impairments when accessing information or navigating environments. They enable equal access to laboratories, museums, historical sites, and other educational settings, allowing for independent and hands-on learning. The virtual nature of the metaverse removes physical barriers, opening access to diverse learning content. This creates opportunities for inclusive, high-quality education and lifelong learning—not only for people with visual impairments, but also for older adults and individuals with limited mobility.

## 2.3 Future work

First, we plan to analyze what devices visually impaired people use in which environments and how they utilize them. This will help us understand how they use their devices and the practical inconveniences they encounter, and will help us identify specific requirements for future visually impaired UI/UX design.

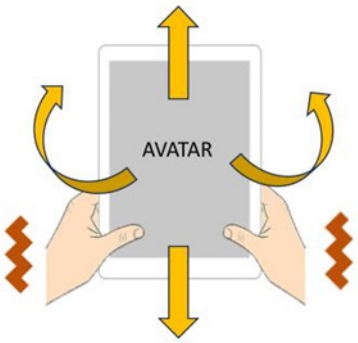
Second, we plan to analyze in detail the Operating System (OS) types and specifications of digital devices such as Personal Computers (PCs), phones, and pads that are mainly used by visually impaired people. By doing so, we will design optimized UI/UX for each device and explore ways to solve the limitations that arise from specific devices or platforms. We will also compare the differences in accessibility features between various devices and aim to design a common interface that can compensate for them.

Third, we will identify these improvement points and develop customized UI/UX enhancements for the visually impaired. Specifically, we will implement innovative technical solutions that reflect the characteristics of each device and user feedback to substantially improve accessibility. This will create a technical standard that can be utilized consistently across different environments.

# 3 Use Case Requirements

- **REQ-01**: Touch UIs for visually impaired users should be designed for consistent usability, with key function buttons large and clearly placed. Each button should contain only essential functions and provide a simplified layout that users find intuitive.
- **REQ-02**: The tactile display should work reliably via Bluetooth and provide real-time feedback to the user through tactile sensations.
- **REQ-03**: Avatar movement should be easy and simple to control by tilting the device, allowing for free navigation within the Metaverse, even in the presence of physical obstacles or complex structures.
- **REQ-04**: The "Avatar Radar" feature should detect events, assets, or other users (avatars) in the user's vicinity and communicate this through voice and alarm feedback.
- **REQ-05**: The device and UI should be compatible across different OS (Windows, Android, iOS, etc.) and device types (PC, Phone, Pad, etc.) and meet accessibility standards.

## 4 Sequence Diagram



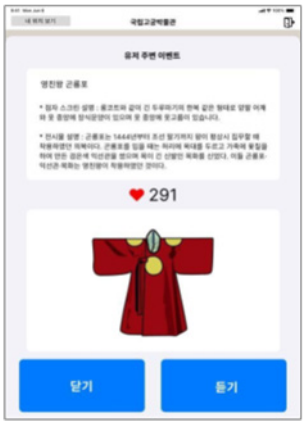



Exhibit Description pop-up Nearby avatar descriptions pop-up

## 5 Reference

[1] eorjs012. Blind_DemoApp_iOS. GitHub. Accessed June 19, 2025. https://github.com/eorjs012/Blind_DemoApp_iOS .

# Use Case 8: AI-Powered Indian Sign Language (ISL) Detection and Translation

**Organization:** International Institute of Information Technology, Hyderabad.

**Country**: India

**Contact Person(S)**:

Pandey, Aishani, aishani.pandey@research.iiit.ac.in

Sachdeva, Arush, arush.sachdeva@research.iiit.ac.in

Mathur, Vivek-vivekofficialwork1@gmail.com

## 1 Use Case Summary Table

| Item | Details |
|---|---|
| Category | Accessibility |
| Problem Addressed | There is a big communication gap between people who speak ISL and those who don't. For the deaf community, social integration and information access are hampered by the absence of real-time translation tools. |
| Key Aspects of Solution | Data Acquisition: Gather and preprocess Indian Sign Language(ISL) datasets from a variety of pertinent and helpful open-source sources, such as expert-curated databases, publicly accessible sign language resources, and YouTube tutorials. To ensure accurate representation, specific context, vocabulary, phrases, and videography protocols related to the ISL context will be applied during data curation.<br>Tech Stack: Python (Pandas, Open Computer Vision (OpenCV)), YouTube Application Programming Interface (API), Web Scraping (BeautifulSoup), Video Annotation Tools (e.g., Visual Geometry Group (VGG) Image Annotator)<br>Data Enhancement: Augment data for low-resource variations using pose estimation, synthetic data generation, and adversarial augmentation techniques to improve generalization.<br>Tech Stack: TensorFlow, PyTorch, OpenPose, Generative Adversarial Networks (GANs) for augmentation |

(continued)

| Item | Details |
|---|---|
| | Model Development: Use either the ensemble of current Convolutional Neural Network (CNN)/Transformer models or by identifying key characteristics from each and creating a new model to implement a lightweight, real-time ISL recognition model that is optimised for efficiency on edge devices, guaranteeing a seamless user experience and low latency. We can use language models to help us preserve the emotion and context. Recent advancements in continuous sign language recognition, like the use of motor attention mechanisms [1] and multi-scale feature enhancement [2], suggest potential improvements in real-time detection. A recent study by Hirooka et al. [3] demonstrates the use of Stack Transformer models with spatial-temporal attention for dynamic multi-culture sign language recognition, which can offer significant improvements in dynamic ISL detection by preserving temporal context and motion features.<br><br>Tech Stack: TensorFlow, PyTorch, OpenCV, CNNs (ResNet, VGG), Transformer Models (Bidirectional Encoder Representations from Transformers(BERT)), Generative Pre-trained Transformer(GPT)), Keras, MobileNet for real-time inference on edge devices<br><br>Intermediate Representation: Extract intermediate representations in latent and embedding space, such as:<br>– Skeletal keypoint sequences for tracking of user.<br>– Encoding for mapping signs to linguistic structures.<br>– Temporal feature representation to capture the timing and transitions between signs using detection of hand movements.<br>– Facial feature embedding to track and preserve emotion and expression (with de-personification to ensure privacy).<br><br>Tech Stack: OpenPose, Dlib (for facial feature extraction), TensorFlow/Keras (for embedding), PyTorch (for dynamic embeddings)<br><br>Privacy-Preserving Mechanisms: Ensure anonymization of data using pose estimation for skeletal key points and avoid using facial or personal identifiers. Encrypt data during transmission to prevent unauthorized access.<br><br>Tech Stack: Cryptography Libraries (PyCryptodome, Open Secure Sockets Layer(SSL)), Secure Data Transmission (Hyper Text Transfer Protocol Secure (HTTPS), Transport Layer Security (TLS)), General Data Protection Regulation(GDPR)-compliant Data Storage Systems<br><br>Model Evaluation and Fine-Tuning: Compare performance between existing Transformer and CNN models and our novel model for ISL recognition using various interpretability techniques and feature importance analysis.<br><br>Tech Stack: Shapley Additive Explanations (SHAP), Local Interpretable Model-agnostic Explanations (LIME) |
| Technology Keywords | AI, Natural Language Processing(NLP), Sign Language Recognition, Pose Estimation, Real-time Translation, Embedding Space |
| Data Availability | Public |
| Metadata (Type of Data) | Video, Text (ISL videos annotated with linguistic mapping) |
| Model Training and Fine-Tuning | CNN, Transformer-based NLP, Pose Estimation, Adversarial Augmentation |
| Testbeds or Pilot Deployments | P2SLR:Privacy-Preserving Sign Language Recognition[4] |

# 2 Use Case Description

## 2.1 Description

### Context and Background

Millions of people in India who are deaf or hard of hearing rely on ISL as their main form of communication. Nonetheless, there is still a big communication gap between people who speak ISL and those who don't. For the deaf community, social integration and information access are hampered by the absence of real-time translation tools. The language barrier that separates ISL from commonly spoken languages like Hindi and English also makes communication more difficult and restricts the participation of deaf people in public services, education, and mainstream conversations. While the core focus of this work is on developing an accurate ISL recognition model, we recognize that effective communication solutions must reflect the rich sociocultural diversity of India's deaf and hard-of-hearing communities. Incorporating detailed user archetypes and narratives that capture lived experiences and regional linguistic diversity will be essential for real-world adoption and accessibility. These qualitative insights will guide future phases of this project to ensure the technology is inclusive and responsive to community-specific communication barriers.

### Objective and Aims

This research aims to develop an AI-powered system that can effectively detect and translate ISL into English, and vice versa. The core objectives include:

- Real-time ISL Recognition and Translation: Enabling real-time translation between ISL and English, allowing smooth interaction between deaf individuals and others.
- Bidirectional Translation: The system will not only translate from ISL to English but also facilitate translation from English to ISL, making it versatile.
- Multi-language Support: Since translation from other languages (e.g., Hindi) to ISL will be enabled by first converting the text to English, this approach will support indirect translation to ISL from a variety of languages.
- Proof-of-Concept (PoC) Development: The project will create a PoC system to demonstrate the feasibility of ISL recognition and translation in real-time using open-domain ISL datasets and lightweight machine learning models. Use models like the CLIP-SLA method [5], which offers parameter efficient adaptations for continuous ISL recognition.

**Use Case Status:** In Development (PoC Stage) and this are currently developing a PoC system.

Partner: N/A

## 2.2 Benefits of the use case

Quality Education: Integrating ISL translation tools into educational settings enhances accessibility for deaf and hard-of-hearing learners. This ensures that individuals who rely on sign language can fully engage with classroom activities, online learning platforms, and academic programs, contributing to a more inclusive and equitable education system.

Industry, Innovation, and Infrastructure: The project showcases a forward-thinking application of artificial intelligence and machine learning to drive social good. By developing technologies such as real-time sign language recognition, pose estimation, and natural language processing, the initiative strengthens digital infrastructure and encourages innovation that supports accessibility for all.

Reduced Inequality: The solution directly empowers deaf and hard-of-hearing communities by enabling greater participation in education, employment, and everyday social interactions. By removing communication barriers, it promotes equal access to opportunities, resources, and services, thereby helping to reduce disparities and foster a more inclusive society.

## 2.3 Future Work

- Pilot Study Completion: Conduct a comprehensive pilot study evaluating initial ISL detection and translation models.
- Prototype Development: Deploy a functional PoC system demonstrating realtime ISL recognition and translation.
- Research Paper Submission: Prepare and submit a research paper detailing the methodology, findings, and potential improvements based on the pilot study.
- Dataset Release: Publish an initial dataset of annotated ISL signs, including pose estimation and space embeddings representations.
- Technical Report: Document findings from early testing phases and propose refinements for future work.
- AI for Good Summit Presentation: Demonstrate prototype performance and preliminary results at the AI for Good Summit 2025.
- Avatar Generation: Investigate avatar generation for ISL translation, creating dynamic sign language avatars for enhanced visual translation.
- Ethical Safeguards and Data Privacy Measures: Implement and document robust privacy protections for sensitive video recordings, including anonymization techniques (e.g., skeletal abstraction and facial de-identification), secure storage using GDPR-compliant frameworks, and user-informed consent procedures. A Data Protection Impact Assessment (DPIA) will be conducted prior to any public data release to ensure ethical standards are met.

# 3 Use Case Requirements

- **REQ-01**: Development of an organic dataset with following:
  - Data collection from native signers, YouTube videos, and expert collaboration.
  - Annotation strategies for regional sign language variations.
  - Quality control measures to ensure real-world usability and linguistic diversity.

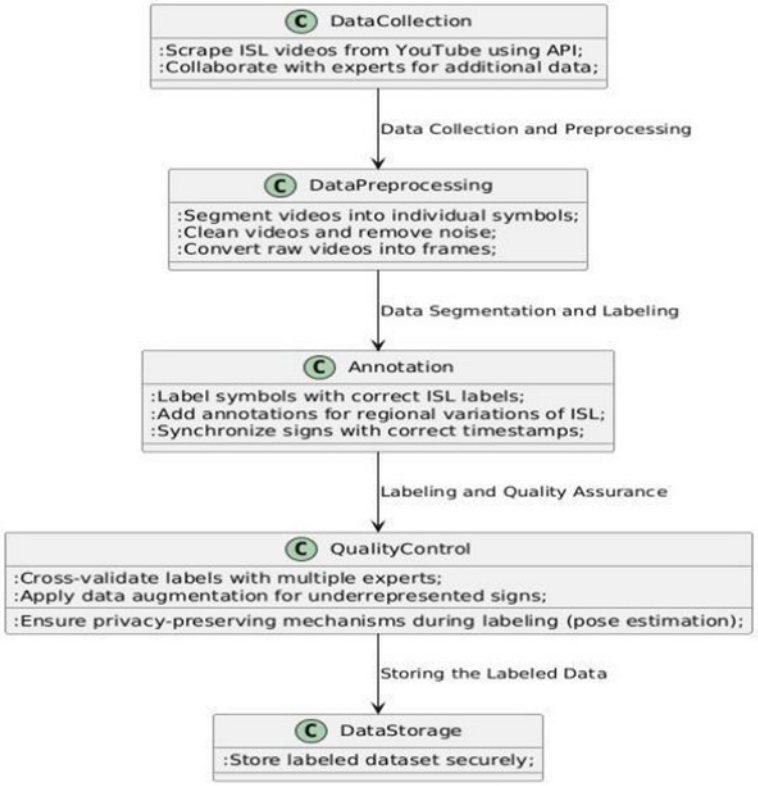
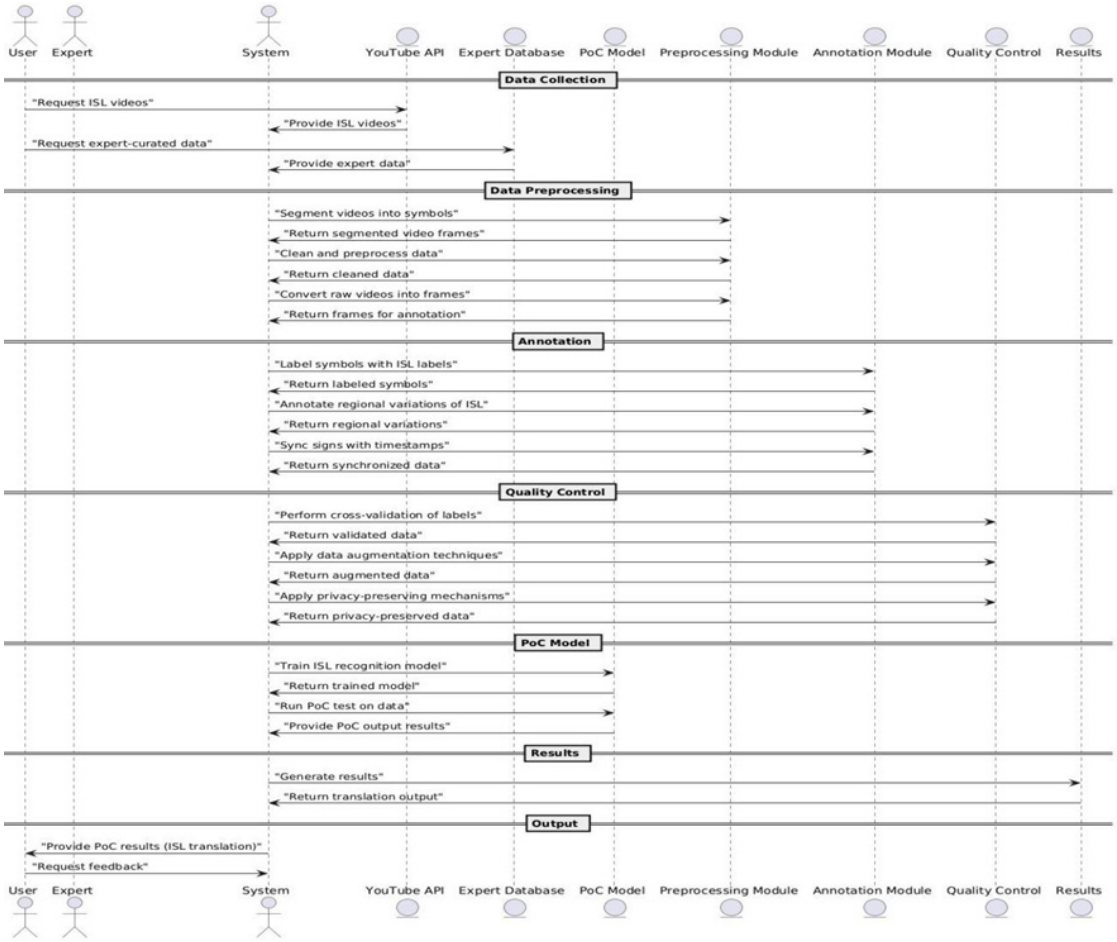



Figure 1: Proposed Data Collection Pipeline

- **REQ-02**: It must handle temporal differences between consecutive signs, including when to update the predicted letter on the screen.
- **REQ-03**: The model should join letters to form words and verify if the word belongs to a dictionary.
- **REQ-04**: The system should understand basic key statement expressions in sign language.
- **REQ-05**: The system should use pose estimation and embedding representations for better accuracy.
- **REQ-06**: The AI model should support translation in local languages such as Hindi and English.
- **REQ-07**: Comparing outputs of existing Transformer and CNN models with our model to understand their performance differences using interpretability techniques, finding feature importance in each of the models.
- **REQ-08**: Develop a pipeline by addressing regional variations, linguistic nuances, and real-world usage scenarios.
- **REQ-09**: Preserve the context and emotion of the message to be conveyed throughout the translation and identification process.

## 4 Sequence Diagram

The following Diagram (High-Level System Architecture for ISL Detection and Translation) Illustrates the proposed System Pipeline:




## 5 References

[1] Q. Zhu, J. Li, F. Yuan, and Q. Gan, "Continuous sign language recognition based on motor attention mechanism and frame-level self-distillation," Machine Vision and Applications, vol. 36, no. 1, pp. 1–12, 2025. doi: 10.1007/s00138-024-01371-5.

[2] Z. Wang, D. Li, R. Jiang, and M. Okumura, "Continuous sign language recognition with multi-scale spatial-temporal feature enhancement," IEEE Access, 2025. doi: 10.1109/ACCESS.2025.1234567.

[3] K. Hirooka, A. S. M. Miah, T. Murakami, Y. Akiba, Y. S. Hwang, and J. Shin, "Stack transformer based spatial-temporal attention model for dynamic multi-culture sign language recognition," arXiv preprint, 2025. eprint: arXiv:2503.16855. [Online]. Available: https://arxiv.org/abs/2503.16855.

[4] V. K. Tanwar, g. sharma gaurav, B. Raman, and R. Bhargava, "P2slr: A privacy-preserving sign language recognition as-a-cloud service using deep learning for encrypted gestures," Feb. 2022. doi: 10.36227/ techrxiv.19064063.v1. [Online]. Available: http://dx.doi.org/10.36227/techrxiv.19064063.v1.

[5] S. Alyami and H. Luqman, "Clip-sla: Parameter-efficient clip adaptation for continuous sign language recognition," arXiv preprint, 2025. eprint: arXiv:2504.01666. [Online]. Available: https://arxiv. org/abs/2504.01666.

# Use Case 9: Intelligent and Automatic System Interface Control System for Users with Disabilities

**Organization:** Addis Ababa Science and Technology University

**Country:** Ethiopia

**Contact Person:**

Mr. Kibremoges Fenta, kibremogesf@gmail.com

## 1 Use Case Summary Table

| Item | Details |
|---|---|
| Category | Accessibility |
| Problem Addressed | Millions of people suffer from physical disabilities such as paraplegia or hand impairments, making it difficult or impossible to interact with digital systems that rely on hand input. As of 2025, the prevalence of hand disability, particularly related to pain and functional limitations, is around 13.6% [1]. Above 650 million people worldwide live with disabilities, including those who struggle to use computers, access the internet, or control software due to the absence of functional hands. Existing systems lack intuitive, hands-free interaction methods. We can address this urgent problem by developing an "Intelligent and Automatic System Interface Control System for Users with Disabilities", which enables hand-disabled individuals to use computers and digital devices without relying on a mouse or touch input, improving accessibility, independence, and digital inclusion. |

(continued)

| Item | Details |
|---|---|
| Key Aspects of Solution | Face and Eye-Based System Control:<br>A smart system that detects and tracks human face and eye movements to interact with and control any digital device, enabling hands-free accessibility for users with physical disabilities.<br>Search Engine Optimization (SEO):<br>Eye-tracking data can be leveraged to analyze user attention and behavior, improving website layout, content placement, and overall SEO strategies.<br>Market Research and Advertising:<br>Eye movement and look tracking offer valuable insights into user engagement, helping businesses optimize advertisements, product placement, and User Experience (UX) design based on real user interaction patterns. |
| Technology Keywords | Artificial Intelligence (AI), Computer Vision, YOLOv5, Dlib, Facial Landmark Detection, Histogram of Oriented Gradients (HOG), Eye Aspect Ratio (EAR), Haar Cascades |
| Data Availability | Some Private and Public [2] |
| Metadata (Type of Data) | Images |
| Model Training and Fine-Tuning | The system uses a hybrid approach combining YOLOv5 and Dlib to detect and track facial features for hands-free interface control. YOLOv5 is trained to detect the face and eyes in real time from video frames at 30 FPS. Once the face region is detected, Dlib extracts 68 facial landmarks, enabling accurate tracking of eye movements, blinks, and head orientation. Blink detection is performed by calculating the Eye Aspect Ratio (EAR) from these landmarks. Finally, using Haar classification algorithms, the detected face and eye gestures are mapped to mouse actions (e.g., cursor movement and clicks), enabling seamless operation of digital systems without hand input. |
| Testbeds or Pilot Deployments | YOLOv5 for facial and eye detection, Dlib for facial landmark extraction, and Eye Aspect Ratio (EAR). [3] |

# 2 Use Case Description

## 2.1 Description

This use case proposes a solution for those who have disabilities to use a computer system with ease, regardless of their difficulties using traditional computers. We have used the various built-in sensors like camera, mic, and touch screen together with keyboard and mouse as input to the system to capture user interactions and to provide the desired output. To accomplish this, our system will be trained using data that consistently updates its model for a more efficient result. The advances in AI have made many tasks that only humans were capable of performing to be done by intelligent agent systems. Through a previous research study, it was concluded that eyes are an excellent candidate for ubiquitous computing since they respond well during interaction with computer systems. Using this underlying information from eye movements could allow bringing the use of computers back to such users. For this purpose, we propose an eye gesture control system that is operated by human eyes only. The purpose of this work is

to design an open-source source generic eye gesture control system that can effectively track eye movements and enable the user to perform actions mapped to specific eye movements/ gestures by using sensors like a camera and a mic as input. It detects the pupil from the user's face and then tracks its movements and gives the desired output, like moving the cursor to the desired location, scrolling, zooming in, zooming out, clicking, etc. Most importantly, it needs to be accurate in real time so that the user is able to use it for everyday life. That will be done by making the AI model consistently update its model by getting data from edge devices to the cloud for continuous improvements of the intended human-computer interaction.

**Use case status:** The use case is part of a project we are working on already.

Partner: Center for Disability and Development in Ethiopia

## 2.2 Future work

Data collection and analysis for more general systems: Additional datasets of facial expressions and eye tracking are required to supplement existing training data, especially considering the availability of facial data to make it inclusive for all users with less bias. Collaborating with healthcare and research institutions for further implementation will be made.

Model development for a general computer-human interaction system: Continuous refinement and optimization of machine learning models for facial expression and eye tracking analysis for all the systems that need physical interaction. This involves training the models on larger and more diverse datasets to enhance accuracy, applicability, functions, and robustness. Iterative model development will enable the system to effectively interpret and respond as a desired output.

Search engine optimization (SEO): SEO for websites and businesses can be done effectively by employing eye-tracking-based information and data analytic. This potential area of research is being spearheaded by Microsoft and Google.

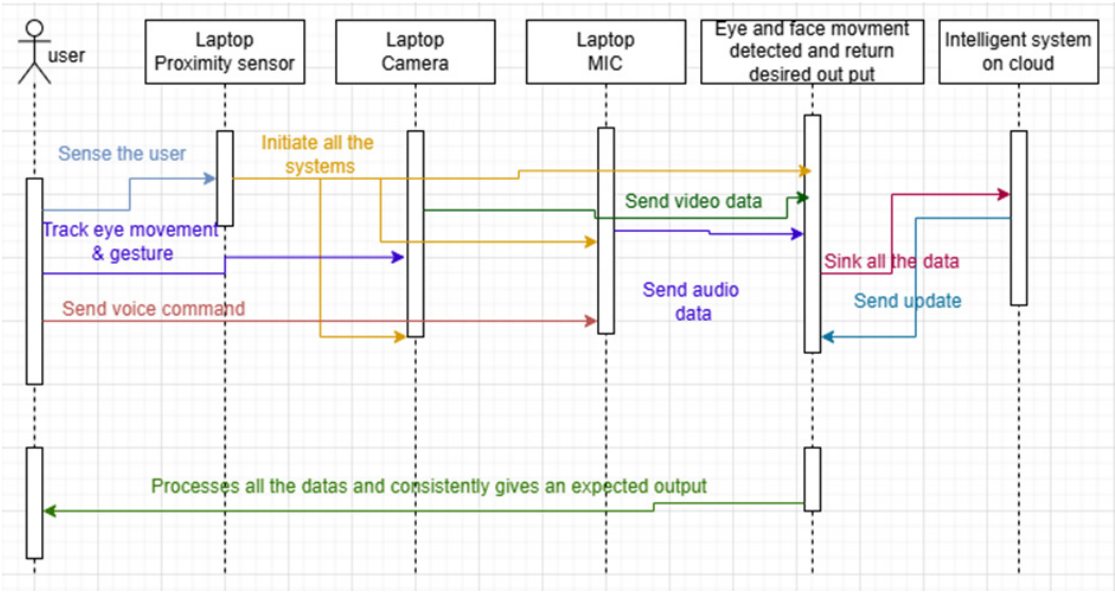
Market research and advertising testing: Face and Eye tracker can play a major role in marketing. To understand the demands of the customer, eye trackers are placed in markers, and they determine for how much time the user is staring at the product. This will help manufacturers and producers to make products that meet consumer requirements.

# 3 Use Case Requirements

- **REQ-01**: It is critical that the system enables individuals with hand disabilities to control digital devices using only face and eye gestures in real time.
- **REQ-02**: It is expected that the system utilizes accurate face and eye detection algorithms (YOLOv5 and Dlib) operating at a minimum of 30 Frames Per Second (FPS ) to ensure smooth user interaction.
- **REQ-03**: It is critical that the system supports blink-based input (e.g., clicks) through EAR analysis from Dlib facial landmarks.
- **REQ-04**: It is critical that the system runs efficiently on low-cost, widely available hardware such as laptops or Raspberry Pi with webcam support.
- **REQ-05**: It is of added value that the system maintains robust performance across diverse lighting conditions, skin tones, and facial structures.

- **REQ-06**: It is expected that the system functions locally without requiring an internet connection to ensure user privacy and security.
- **REQ-07**: It is of added value that the system allows users to customize gesture mappings based on their personal preferences and abilities.
- **REQ-08**: It is critical that the interface remains user-friendly and easily operable by individuals with limited mobility, requiring minimal setup and training.

## 4 Sequence Diagram




## 5 Refereneces

[1] Dahaghin, S., Bierma-Zeinstra, S., Reijman, M., Pols, H., Hazes, J., & Koes, B. (2004). Prevalence and determinants of one month hand pain and hand related disability in the elderly (Rotterdam study). *Annals of the Rheumatic Diseases*, *64*(1), 99. https://doi.org/10.1136/ard.2003.017087

[2] https://www.kaggle.com/datasets/ildaron/dataset-eyetracking: dataset-eyetracking

[3] https://github.com/kbsha/drive/tree/main: Intelligent and Automatic System Interface Control System for Users with Disabilities person

# Use Case 10: Trustworthy Artificial Intelligence for Autonomous Wheelchairs

**Organization:** University of Applied Sciences and Arts of Southern Switzerland (SUPSI), Dalle Molle Institute for Artificial Intelligence (IDSIA), Trustworthy Autonomous Systems (TAS) research group

**Country:** Switzerland

**Contact Person(S)**:

Primary contact: Prof. Francesco Flammini, francesco.flammini@supsi.ch

Secondary contact: Dr. Franca Corradini, franca.corradini@supsi.ch

Any other contacts: Dr. Alessandro Antonucci, alessandro.antonucci@supsi.ch,

Jerome Guzzi, jerome.guzzi@supsi.ch,

Carlo Grigioni, carlo.grigioni@supsi.ch

## 1 Use Case Summary Table

| Item | Details |
|---|---|
| Category | Accessibility |
| Problem Addressed | Development of trustworthy artificial intelligence for safe autonomous wheelchairs supported by drones, with a proof-of-concept of robust perception in a road-crossing scenario |
| Key Aspects of Solution | Development of a reference scenario and related dataset in IDSIA Autonomous Robotics lab (https://idsia-robotics.github.io/) using RoboMasters and ROS for performance evaluation of multi-sensor (wheelchair camera, drone camera, distance sensors, etc.) information fusion compared to the ground-truth generated by the OptiTrack system. Fine tuning performed through usage of pre-trained YOLO networks as well as custom re-trained models for performance evaluation and comparison. Artificial vision robustness evaluation through the application of software filters to simulate adverse environmental conditions in outdoor environments. Performance optimization and fine-tuning of sensor redundancy and diversity. Methodology for probabilistic safety evaluation and self-adaptation. |
| Technology Keywords | wheelchairs, drones, safe autonomy, robust perception, anomaly detection, trustworthy artificial intelligence |
| Data Availability | Public |
| Metadata (Type of Data) | Visual (video recordings), text (incl. Python code) |

(continued)

| Item | Details |
|---|---|
| Model Training and Fine-Tuning | Artificial vision developed and fine-tuned using the latest YOLO versions, using pre-trained networks and custom re-training for performance evaluation and comparision. Training and testing of the models in IDSIA autonomous robotics lab using RoboMasters, ROS, and OptiTrack system. Multi-sensor information fusion schemes tested in different configurations with diverse sensors and different perspectives on the scene. Probabilistic evaluation using Dynamic Bayesian Networks to support safety assessment and dynamic reconfiguration through self-adaptation according to the MAPE-K (Monitor-Analyse-Plan-Execute over a shared Knowledge) autonomic computing paradigm. |
| Testbeds or Pilot Deployments | [9] [10] |
| Code repositories | [11] [12] |

# 2 Use Case Description

## 2.1 Description

In this use-case, we address the potential of trustworthy artificial intelligence for autonomous wheelchairs supported by flying drones. The system is aimed to provide an innovative s olution for enhancing mobility and independence of motion-impaired people in future smart-cities. Safety is a critical concern when deploying such systems, therefore we focus on how to achieve safe autonomy in real-world scenarios such as road crossing, which is used as a proof-of-concept. To that aim, we leverage on diversity and redundancy of sensing technologies together with self-adaptation through probabilistic models to capture uncertainty and non-stationarity in the environment. The approach has been developed within a European project named REXASI-PRO (REliable and eXplAinable Swarm Intelligence for People with Reduced mObility)[1].

The project focuses on autonomous wheelchairs that can significantly improve the quality of life of people with reduced mobility and sensory impairment by enabling safe door-to-door navigation.

Autonomous wheelchairs are not currently available on the market. The visionary concept of combining wheelchairs with drones has the potential to improve the sensing capabilities and robustness for autonomous navigation through environmental mapping, enhanced obstacle detection, and danger estimation. Those are performed by leveraging on multi-sensor information fusion, deep learning, and artificial vision. Another innovative aspect is represented by the focus on trustworthy artificial intelligence (AI), to ensure that the autonomous operation of the wheelchair will be compliant with relevant norms and regulations, and it will be robust and explainable to be certified against international safety standards, which is currently a grand challenge for all classes of autonomous vehicles. We focused on a reference scenario of road-crossing to provide a proof of concept of the approach (see Figure 1). The final objective is to enable run-time sensor reconfiguration through self-adaptation and to demonstrate compliance to quantitative safety integrity requirements through Dynamic Bayesian Networks (DBN) (see references [4] and [3]).

By focusing on people with reduced mobility and sensory impairment, the project addresses ethical principles of social inclusiveness and non-discrimination. Being an open project, all its results, deliverables, and datasets are public, so that they can be used and further developed by the research community worldwide.

The Swiss part of the project was funded by SERI under contract 22.0029. The global project has been selected within the Horizon Europe R&I programme under grant agreement 101070028 (call HORIZON-CL4-2021-HUMAN-01-01).

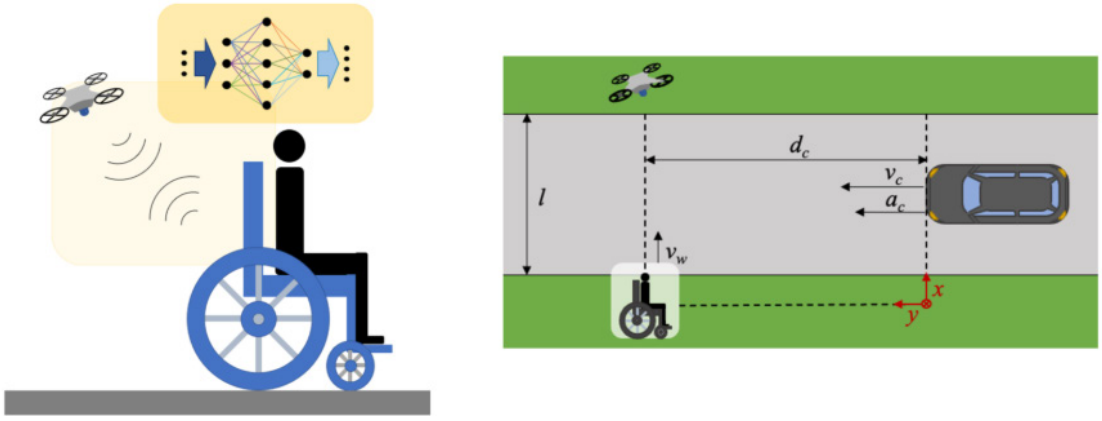



**Figure 1: The wheelchair-drone concept (left) and the reference road-crossing scenario (right).**

### Use Case Status:

We developed a laboratory demonstrator and specific dataset addressing outdoor navigation scenarios, such as road-crossing, to evaluate and improve the safety of autonomous wheelchairs in real-world environments. For the road-crossing scenario, we developed an analytic danger function to be evaluated at run-time using information from diverse sensors and compared the results in different hardware and software configurations, including pre-trained and custom YOLO models (see reference [2]). We benchmarked diversely trained YOLO models for artificial vision in combination with other distance sensors (i.e., lidars in the final prototype) both in nominal and degraded operating conditions, i.e., in presence of environmental disturbances such as adverse weather conditions, to test perception robustness (see reference [1]).

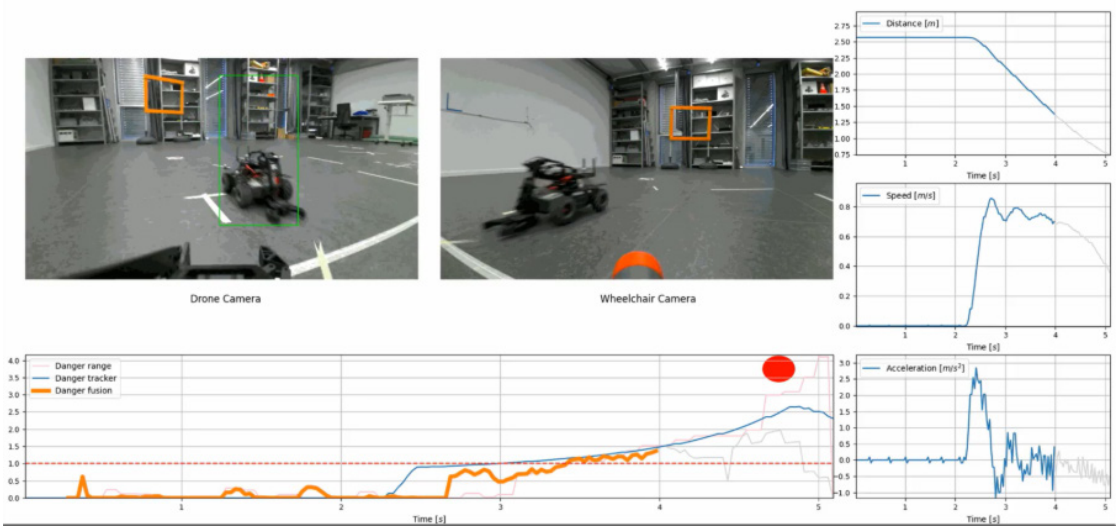



Figure 2: Experimental evaluation of road-crossing danger function in IDSIA Autonomous Robotics Lab.

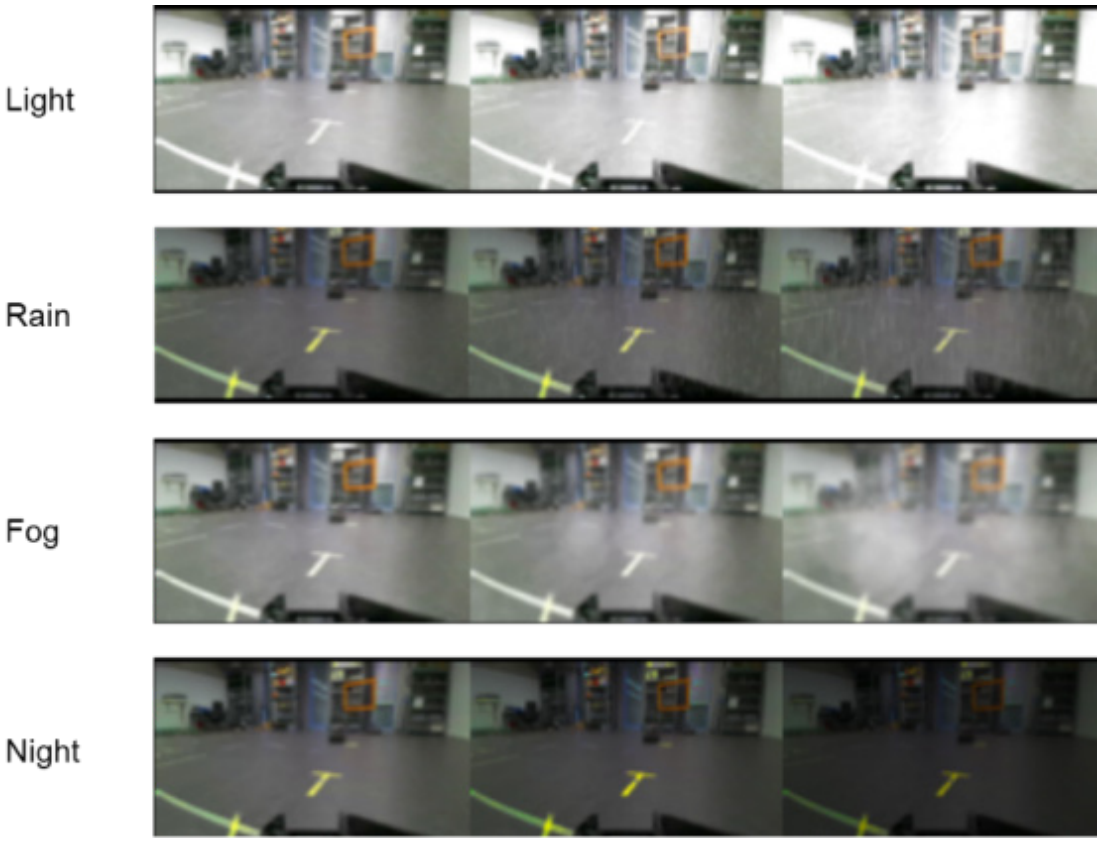



Figure 3: Application of software filters to simulate adverse environmental conditions.

**Partners:**

Spindox Lab, EU project coordinator

By coordinating the EU project, Spindox Lab ensures integration of different contributions from project partners and demonstration using real-world prototypes.

EU Project Partners (V-Research, Aitek, CNR, University of Seville, Euronet, Hovering Solutions, King's College London, DFKI AI)

The other EU project partners develop specific wheelchair and drone technologies for the prototypes to be used in final demonstration.

## 2.2 Benefits of the use case

The project focuses on safe autonomous wheelchairs that can significantly improve the quality of life of people with reduced mobility and sensory impairment by enabling safe door-to-door navigation. The system is aimed to provide an innovative solution for enhancing mobility and independence of motion-impaired people in future smart-cities.

By focusing on people with reduced mobility and sensory impairment, the project addresses ethical principles of social inclusiveness and non-discrimination.

## 2.3 Future work

The research team is currently working in several directions.

1. Improve accuracy in event detection by fine tuning YOLO-based artificial vision through diversity at software-level (using diverse networks and training data) and perspective-level (using diverse points of view from wheelchair and drone)
2. Improve robustness to disturbances by adaptive multi-sensor integration and reputation management at different data-fusion and decision-fusion levels
3. Develop additional operational scenarios such as the ongoing experimentation with surface anomaly detection to prevent accidents due to uneven terrain and hazardous floor conditions.

Research data, including training and testing datasets, will be made available on public repositories.

Final project demonstration using real prototypes is expected before funded project ending in Fall 2025. However, research is planned to continue after that date. A total of three senior researcher and three junior researchers (one visiting) at IDSIA are currently working on the project. Among those, PhD student Franca Corradini is developing her research on this specific use-case, cooperating with other PhD students and AI researchers from other universities (e.g., University of Salerno, Italy), and she is only at the middle of her journey to getting her doctoral degree.

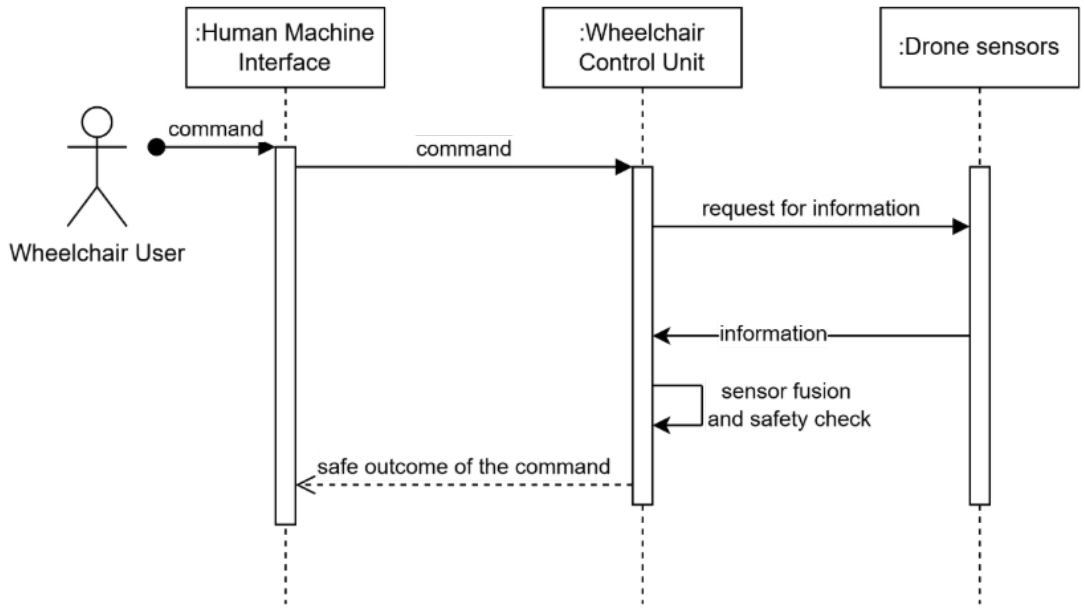
## 3 Use Case Requirements

The following requirements only address the perception subsystem of the wheelchair-drone system. Further requirements are needed to manage the overall navigation.

- **REQ-01**: The wheelchair-drone system must be able to detect and evaluate threats, dangers, and anomalies in the surrounding environment both in indoor and outdoor scenarios, with high accuracy using redundant and diverse sensing technologies.
- **REQ-02**: The wheelchair-drone perception system must be robust against sensor faults and environmental disturbances such as adverse weather conditions (rain, fog, etc.), and adapt its sensing configuration accordingly.
- **REQ-03**: The wheelchair-drone system must be able to take safe autonomous decisions in critical scenarios such as road crossing.
- **REQ-04**: The safety of the wheelchair-drone perception system must be quantitatively evaluated at both design-time and run-time to enable dynamic assessment, self-adaptation, and to support certification against relevant standards.
- **REQ-05**: For safety evaluation, explainable and interpretable probabilistic models must be used to account for sensing accuracy, fault-modes, and environmental uncertainty.
- **REQ-06**: In order to efficiently manage computational workload when running machine learning algorithms, the wheelchair must be equipped with a control unit of sufficient video processing, data storage, and communication capabilities to guarantee fast and predictable response times in real-world operational scenarios.
- **REQ-07**: For real-time fail-safe operation, due to possible latencies, communication between wheelchair and drone must use timestamps to detect excessive response delays and react accordingly, e.g., by slowing down or even stopping the wheelchair.

## 4 Sequence Diagram

The sequence diagram shows an abstract scenario where the wheelchair user issues a command that is managed by the wheelchair control unit interrogating distributed sensors to increase perception safety.



## 5 Refereneces


[1] Corradini, F., C. Grigioni, A. Antonucci, J. Guzzi, and F. Flammini. 2024. "Experimental Evaluation of Road-Crossing Decisions by Autonomous Wheelchairs against Environmental Factors." TechRxiv, June 3, 2024. https://doi.org/10.36227/techrxiv.171742519.93192238 CloudFront

[2] Grigioni, C., F. Corradini, A. Antonucci, J. Guzzi, and F. Flammini. 2024. "Safe Road-Crossing by Autonomous Wheelchairs: A Novel Dataset and Its Experimental Evaluation." In Computer Safety, Reliability, and Security. SAFECOMP 2024 Workshops, edited by A. Ceccarelli, M. Trapp, A. Bondavalli, E. Schoitsch, B. Gallina, and F. Bitsch, Lecture Notes in Computer Science, vol. 14989. Cham: Springer. https://doi.org/10.1007/978-3-031-68738-9_4 Springer LinkarXiv

[3] Corradini, F., F. Flammini, and A. Antonucci. 2023. "Probabilistic Modelling for Trustworthy Artificial Intelligence in Drone-Supported Autonomous Wheelchairs." In Proceedings of the First International Symposium on Trustworthy Autonomous Systems (TAS '23). New York: Association for Computing Machinery. Article 52, 1–5. https://doi.org/10.1145/3597512.3599716 arXiv

[4] Flammini, F., S. Marrone, R. Nardone, M. Caporuscio, and M. D'Angelo. 2020. "Safety Integrity through Self-adaptation for Multi-sensor Event Detection: Methodology and Case-study." Future Generation Computer Systems 112: 965–981. https://doi.org/10.1016/j.future.2020.06.036

[5] SUPSI. 2024. "Use-Case: Autonomous Wheelchair Support." Accessed August 25, 2025. https://www.supsi.ch/en/l-ai-per-il-supporto-di-persone-con-disabilit%C3%A0

[6] SwissInfo. 2024. "Swiss Drone-Assisted Wheelchair Showcases AI's Promise and Problems." Accessed August 25, 2025. https://www.swissinfo.ch/eng/science/swiss-drone-assisted-wheelchair-showcases-ais-promise-and-problems/87686116

[7] Euronews. 2024. "This Self-driving Wheelchair May Help People with Disabilities Cross the Road." November 4, 2024. https://www.euronews.com/health/2024/11/04/this-self-driving-wheelchair-may-help-people-with-disabilities-cross-the-road

[8] USI. 2024. "Dalle Molle Award Usage." Accessed August 25, 2025. https://www.inf.usi.ch/en/feeds/11002

[9] IDSIA (TAS). 2024. "Safe Road-Crossing by Autonomous Wheelchairs." TAS @ IDSIA, USI-SUPSI blog, March 15, 2024. https://tas.idsia.ch/2024/03/15/safe-road-crossing-by-autonomous-wheelchairs/ tas.idsia.ch

[10] REXASI-PRO. 2024. "Safe Road-Crossing by Autonomous Wheelchairs" (information fusion overview). REXASI-PRO blog. https://rexasi-pro.spindoxlabs.com/safe-road-crossing/ rexasi-pro.spindoxlabs.com

[11] REXASI-PRO. 2024. "Use-Case Dataset on Hugging Face." REXASI-PRO documentation. https://huggingface.co/datasets/carlogrigioni/safe-road-crossing-aw-dataset Hugging Face

[12] Grigioni, C., F. Corradini, A. Antonucci, J. Guzzi, and F. Flammini. 2024. Safe Road-Crossing AW Dataset. GitHub repository. https://github.com/CarloGrigioni/safe_roadcrossing_aw

# 4.10. Category 10: Intelligent Transport

## Use Case 1: Tanbeeh: Revolutionizing Safety and Efficiency with AI

**Country:** United Arab Emirates

**Organization:** DP World

**Contact Person(s)**:

**Primary Contact:** Sabeen Muhammad – Sabeen.Muhammad@dpworld.com, +971 4 8898330, +971-55-2794335

**Secondary Contact:** Marwa Iqbal Mir Khoory – Marwa.Khoory@dpworld.com, +971-4-8896789

### 1 Use Case Summary Table

| Item | Details |
|---|---|
| Category | Industrial Safety<br>Intelligent Transport<br>AI-Driven Security & Monitoring |
| Problem Addressed | Large-scale industrial facilities (e.g., ports and logistics hubs) rely on traditional, manual safety monitoring approaches that are slow and error-prone. This results in delayed fire detection, inefficient PPE compliance checks, bottlenecks in access control, and suboptimal traffic management, all of which increase risks and operational costs. |
| Key Aspects of Solution | • **AI-Powered Computer Vision:** Uses real-time enterprise-grade video processing to detect smoke, fire, and anomalies in large open space environments.<br>• **Automated PPE Compliance:** Vision-based monitoring for ensuring 100% compliance with safety gear requirements.<br>• **Facial Recognition for Access Control:** Streamlines external truck entry by quickly registering/authenticating individuals, reducing gate delays.<br>• **Traffic Analysis:** Provides real-time analytics to predict congestion, optimize routes, and lower carbon emissions in a bounded freezone (57 sq km) that is internally managed by DP World.<br>• **Integration:** Seamlessly works with existing CCTV and IT infrastructure for rapid deployment. |
| Technology Keywords | Artificial Intelligence, Computer Vision, Neural Networks, YOLOv8, Deep Learning, Real-Time Monitoring, Facial Recognition, Traffic Analytics, PPE Compliance |

(continued)

| Item | Details |
|---|---|
| Data Availability | • **Availability:** public and privately held data (proprietary datasets and operational camera feeds)<br>• **Sources:** Internal DP World data, including feeds from hundreds of CCTV cameras. And public datasets |
| Metadata (Type of Data) | • Images and video streams<br>• log data<br>• Analytical reports and alert notifications |
| Model Training and Fine-Tuning | Tanbeeh uses customized YOLOv8 models with approximately 750 million parameters. The system was trained on 100,000 images collected from diverse environments (factories, ports, offices, etc.) with 500 hours of GPU training, and is continuously updated using live data feedback to refine detection accuracy. |
| Testbeds or Pilot Deployments | • Deployed live at DP World since November 1, 2023<br>• Currently connected to 125 CCTV cameras with plans to scale up<br>• Actively used across various departments (HSE, Traffic &<br>• Security, Operations, Crisis Management) |
| Code repositories | While the code remains proprietary, an overview and further technical details can be accessed via public information at [1]. |

# 2 Use Case Description

## 2.1 Description

Tanbeeh is an AI-driven safety and security platform specifically developed for industrial environments such as DP World's port facilities. It leverages advanced computer vision, real-time data analytics, and deep learning to automate:

- **Fire & Smoke Detection:** Reducing traditional detection times from 25–30 minutes to mere seconds (currently 1 second)
- **PPE Compliance Monitoring:** Automatically verifying safety gear compliance, eliminating labor-intensive manual checks, PPE compliance accuracy rose to 70% (from 50%) under the AI-based SMART detection system, fully eliminating the error-prone manual checks
- **Facial Recognition:** Securing access control and reducing gate delays by rapidly authenticating personnel, Facial recognition cuts gate-entry turnaround from 3 minutes down to 15 s.
- **Traffic Analysis:** Monitoring and optimizing traffic flow to reduce congestion and emissions.

**Data Sources:**

Tanbeeh connects to various cameras and pulls frames at FPS (Frames Per Second). Each module might have a different FPS, A grabber process fetches live frames from cameras and sends them to respective AI models for inference.

Since inference is centralized within a very fast LAN (5 Gbps), real-time analytics are possible, and various dashboards show real-time information.

### Data Privacy:

Tanbeeh is fully deployed on-prem in a secure network, no data is shared on the cloud, data privacy is guaranteed, and fully aligned with DP World security policies.

Tanbeeh integrates with other systems within the DPW vicinity using secured APIs.

By replacing manual oversight with automated detection, the system improves safety protocols, boosts operational efficiency, and enhances crisis management capabilities. This case is part of a larger product development initiative that continuously evolves with additional features and expansion plans.

### Partners: Primary Partners:

- **DP World HSE Department:** Implements and monitors safety and environmental protocols using Tanbeeh's AI capabilities.
- **DP World Traffic & Security Department:** Utilizes the platform for facial recognition at entry points and advanced traffic management.
- **DP World Operations and Crisis Management Teams:** Enhance overall operational efficiency and respond swiftly during emergencies.
- **JAFZA (Jebel Ali Free Zone):** Collaborates to ensure seamless logistics and service delivery within the free zone.

### Additional Potential Partners:

- Government entities (e.g., transport authorities, civil defense)
- Security service providers
- Smart city management organizations

### Collaboration:

We had a collaboration with Khalifa University to improve the accuracy of Smoke & Fire models, and we showed a case to their students ways of fine tuning the model to fit specific environments.

We also trained several students from various universities on image labeling, which is currently used to build the data sets in Tanbeeh.

Lessons learned from production deployment:

1. Deployment was extremely challenging, especially when scaling up to connect to 100s of cameras.
2. MLops requires experience in optimizing AI models through pruning and quantization. This must be done carefully, as it might result in reducing accuracy.
3. Acquiring hardware is difficult due to GPUs procurement being slow due to US restrictions.

### Impact on Intelligent Transport and Environmental Safety:

Tanbeeh leverages AI-driven computer vision to transform safety, security, and operational efficiency within industrial environments. By automating fire detection, PPE compliance, and traffic management, it strengthens infrastructure resilience and supports safer, smarter operations at large-scale logistics hubs like DP World.

Its AI-powered traffic analytics optimizes logistics and resource allocation, significantly improving industrial efficiency. By reducing congestion and enhancing traffic flow, Tanbeeh contributes to smarter urban mobility systems. Real-time traffic monitoring minimizes delays and enables better coordination between logistics centers and city infrastructure, facilitating seamless transport operations.

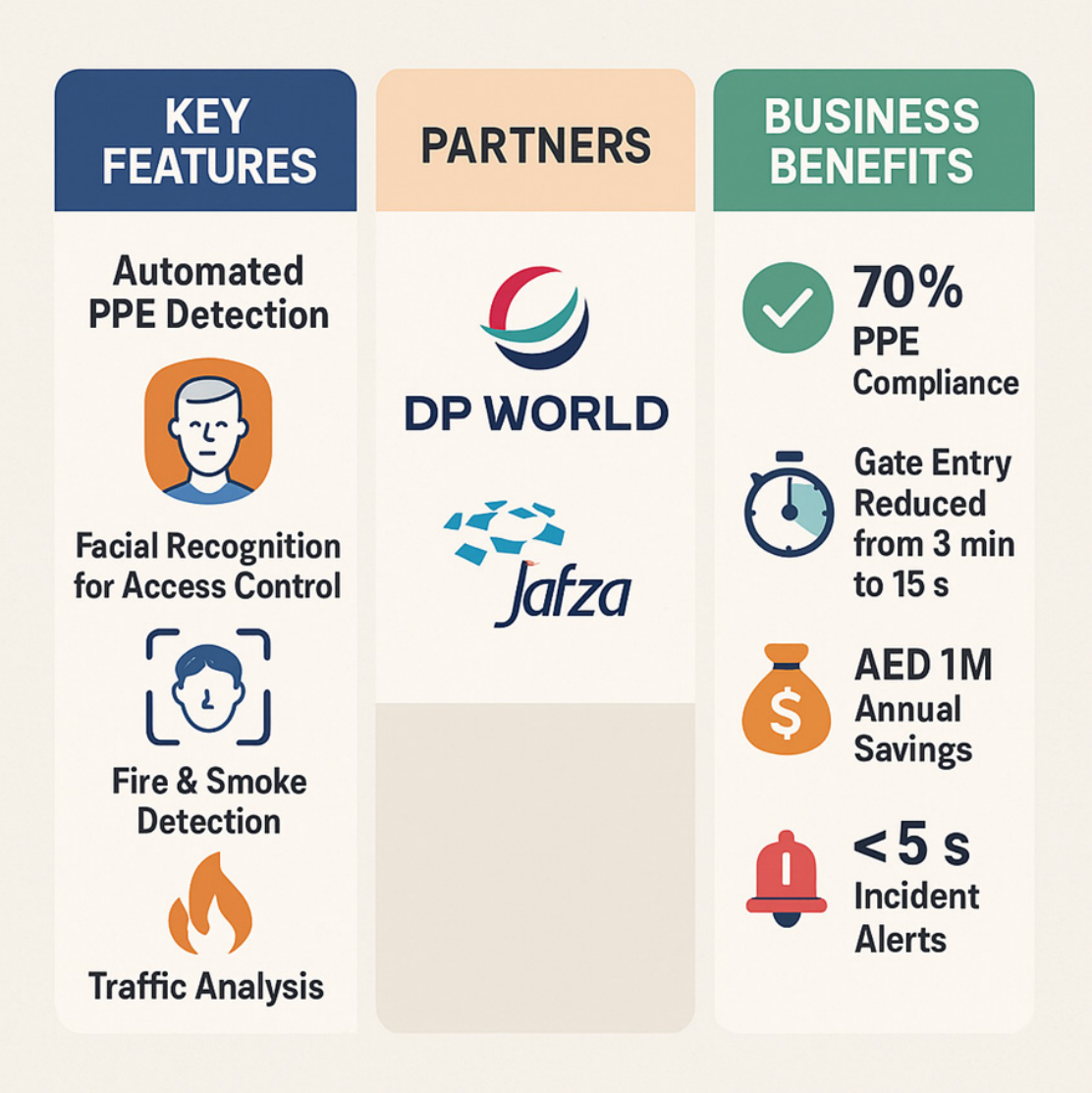
Traffic congestion is a key source of carbon emissions. Tanbeeh's AI-driven optimization reduces vehicle idle times, lowering greenhouse gas emissions and improving air quality in urban and industrial areas. Additionally, early detection and prevention of industrial fires reduce air pollution and toxic emissions, supporting cleaner environments and public health.

Beyond industrial applications, Tanbeeh's early smoke and fire detection capabilities can be extended to forest conservation efforts. By identifying potential wildfires before they spread, the system helps protect biodiversity, ecosystems, and natural habitats, preserving wildlife and preventing environmental damage.

Overall, Tanbeeh integrates advanced AI technologies to enhance intelligent transport, promote environmental protection, and safeguard public safety, creating a foundation for more resilient, sustainable urban and industrial landscapes.

## 2.2. Use case benefits:

- **Eliminates human error in PPE audits:** Automated vision-based PPE detection delivers 70% gear compliance, removing manual-check inaccuracies and streamlining safety audit processes.
- **Drastically improves incident response time:** AI-driven fire & smoke detection slashes alerting delays from 25–30 min to <5 s, accelerating emergency actions and minimizing potential downtime.
- **Reduces operational costs:** Automating safety monitoring and access control yields estimated annual savings of ~ AED 1 million by cutting labor and incident-related expenses.
- **Optimizes gate throughput KPI:** Facial recognition accelerates truck-entry processing from 3 min to 15 s, reducing queue times and congestion costs.
- **Lowers unplanned downtime and audit overhead:** Predictive analytics and automated incident reporting proactively identify risks, reducing downtime and enhancing compliance readiness.




## 2.3 Future work

**Future Plans:**

| Work-stream | High-level Deliverables | Status / Time-range* |
|---|---|---|
| New Vision-AI Models | • Advanced Livestock Counting<br>• Yard Utilization (occupancy reports)<br>• Fall Detection | UAT during FY 2025; all models to reach production the same fiscal year except Yard Utilization, which will go live in the subsequent release window. |

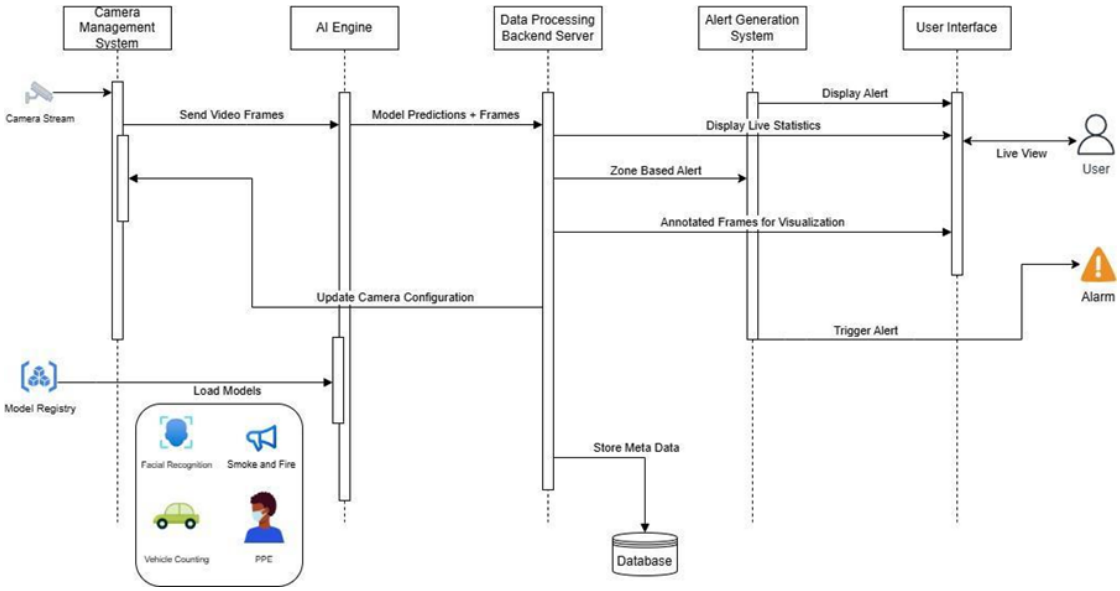
(continued)

| Work-stream | High-level Deliverables | Status / Time-range* |
|---|---|---|
| Enhancements to Existing Models | • Fire & Smoke: accuracy uplift from 79 % -> ≈ 90 % via VL (Llama-4 Scout) integration<br>• Traffic: add more vehicle classes; accuracy boost<br>• PPE: detect gloves, boots, glasses<br>• Facial Recognition: expand registered-driver database | Parallel delivery with new-model work, same FY 2025 window. |
| Platform & User-Experience Upgrades | • Revamped Web UI/UX (modernized dashboards, low-latency image viewer)<br>• Mobile App with extended capabilities (alert triage, live video snapshots, offline mode).<br>• Predictive Analytics & Automated Incident Reporting | Prototype Q3 FY 2025 -> UAT Q3 FY 2025 -> production by year-end. |
| Infrastructure Scale-Up | • On-prem GPU cluster refresh (compute + storage) •Docker / micro-service hardening<br>• CI/CD & monitoring extensions for multi-tenant deployments | Hardware procurement front-loaded to Q1 to unblock model onboarding. |
| Documentation & Enablement | • Ops runbook, model cards, API/SDK bundles<br>• Customer-facing sales collateral and demo environment | Iterative drafts aligned to feature readiness. |

## 3 Use Case Requirements

- **REQ-01**: It is critical to provide real-time video processing capabilities with sub-second inference speeds to ensure timely detection and response in dynamic industrial environments.
- **REQ-02**: It is mandatory to seamlessly integrate with existing CCTV and sensor systems across large-scale industrial facilities, ensuring smooth deployment and leveraging existing infrastructure.
- **REQ-03**: It is critical to ensure high-accuracy detection of events such as fire, smoke, PPE non-compliance, and facial recognition, using custom-trained AI models tailored to the specific industrial context.
- **REQ-04**: It is mandatory to enable system scalability that supports hundreds to thousands of camera feeds simultaneously, ensuring performance remains robust as deployment scales.
- **REQ-05**: It is critical to establish robust data storage, security, and privacy protocols for all operational safety data, protecting sensitive information and ensuring compliance with safety and regulatory standards.
- **REQ-06**: It is mandatory to maintain continuous learning and regular model updates through automated feedback loops, allowing the system to adapt to new scenarios and improve over time.
- **REQ-07**: It is critical to deliver a user-friendly dashboard for real-time monitoring, event reporting, and alert management, enabling operators to respond quickly and effectively to incidents.

## 4 Sequence Diagram




## 5 References

[1] Ashary, E. Bahaudien, et al. "Enhancing Resilience in Next-Generation Wireless Networks Through Deep Learning for Security Enhancement." IEEE Transactions on Consumer Electronics 70, no. 3 (August 2024): 6284–6292. https://doi.org/10.1109/TCE.2024.3415386.

[2] Salhab, W., D. Ameyed, F. Jaafar, and H. Mcheick. "A Systematic Literature Review on AI Safety: Identifying Trends, Challenges, and Future Directions." IEEE Access 12 (2024): 131762–131784. https://doi.org/10.1109/ACCESS.2024.3440647.

[3] Tanbeeh.ai. "Home - Tanbeeh AI." Accessed June 19, 2025. https://www.tanbeeh.ai/.

[4] The National. "Sheikh Mohammed Honours Winners of the UAE AI Award." The National News, November 5, 2024. https://www.thenationalnews.com/news/uae/2024/11/05/sheikh-mohammedhonours-winners-of-the-uae-ai-award/.

[5] World Arabia. "Sheikh Mohammed Honors UAE AI Award Winners." Accessed June 19, 2025. https://world-arabia.com/articles/sheikh-mohammed-honors-uae-ai-awardwinners/.

# Use Case 2: Apollo Go: Pioneering Inclusive Mobility for the Visually Impaired

**Organization**: Baidu

**Country**: China

**Contact Person(s)**:

Email(s): Cathy Fu, fuduanling@baidu.com

Yang Liu, liuyang145@baidu.com

Tianmin Long, intlcomm@baidu.com

## 1 Use Case Summary Table

| Item | Details |
|---|---|
| Category | Accessibility |
| Problem Addressed | The visually impaired community in China faces significant mobility challenges, particularly when navigating urban environments and using public transport. |
| Key Aspects of Solution | The Apollo Go Accessibility Initiative provides tailored solutions to enhance the public transportation accessibility for this group, including Intelligent Arrival Alerts, Zero-Contact Access, Voice-First Control, etc. |
| Technology Keywords | Autonomous Driving Foundation Model (ADFM), Autonomous Vehicles, Artificial Intelligence AI , Machine Learning, Voice Interaction, Bluetooth Proximity, Computer Vision, Safety Redundancy, Large Vision Models, Level 4 Autonomous Driving |
| Data Availability | Private |
| Metadata (Type of Data) | Audio, visual |
| Model Training and Fine-Tuning | AI/Machine Learning(ML) Techniques: The model employs large-scale vision models for real-time environmental perception and navigation, ensuring safe autonomous driving. |
| Testbeds or Pilot Deployments | Accessibility Initiative User Program; Live Fleet Roll-out of Accessibility Care Mode. |

# 2 Use Case Description

## 2.1. Description

China is home to 17 million people with visual impairments, the largest such population in the world[1]. Nearly 50% of them (7.8 million) venture out weekly, yet only around 3 million -less than 18% of the visually impaired population - can navigate without assistance.

In their daily travels, the visually impaired community face multiple challenges including:

- Complex urban environments: Differences in car door handle heights and opening mechanisms make it difficult to locate and open doors accurately.
- In-vehicle inconvenience: Simple tasks like adjusting the air conditioning or playing music often require assistance, while exiting the vehicle poses safety risks due to difficulty detecting approaching vehicles or pedestrians.
- Severe guide dog shortage: With just 200 certified animals for 17 million visually impaired individuals - is compounded by persistent access denials on public transport, leaving even these rare mobility aids underutilized

Committed to its social responsibility of making travel simpler for everyone, Apollo Go, the company's autonomous ride-hailing service platform (robotaxi), has launched the Accessibility Initiative, inviting visually impaired individuals to join a long-term development program for accessibility features[2]. The platform has introduced measures such as free rides for the visually impaired, automatic Bluetooth unlocking, and improved voice interaction features.

Apollo Go has rolled out a special "Accessibility Care" feature tailored to the needs of visually impaired passengers. Key Features of the Accessibility Care Mode:

- Intelligent Arrival Alerts: Upon arriving at designated pickup points, vehicles emit an audible prompt—such as "Apollo Go has arrived"—to help visually impaired passengers locate their ride.
- Zero-Contact Access: Bluetooth proximity sensing enables automatic door unlocking as the passenger approaches, accompanied by a voice greeting that includes the license plate number.
- Voice-First Control: After fastening the seatbelt, passengers can simply say "Let's go" to close the door and begin their journey—no physical interaction required.

The integration of auto-location tracking, hands-free entry, and enhanced voice interaction system further elevated the ride experience. Through real-time audio updates on the surrounding environment, the system helps ensure greater independence and safety for visually impaired passengers, making travel more secure and reassuring.

**Use Case Status:** The use case is in limited public rollout, with Accessibility Care features deployed on a subset of Apollo Go vehicles for visually impaired passengers.

**Partners**: N/A

## 2.2 Benefits of the use case

Industry, Innovation and Infrastructure: Apollo Go's Accessibility Initiative demonstrates how cutting-edge technologies—autonomous driving, Bluetooth proximity sensing, and voice-activated controls—can be woven into transportation systems to deliver truly inclusive services. By

building these capabilities into its core platform and co-creating features with visually impaired users, Apollo Go not only advances mobility innovation but also drives applied research in AI, human-machine interaction, and smart infrastructure with a clear focus on social responsibility.

Sustainable Cities and Communities: By equipping autonomous vehicles with intelligent arrival alerts, zero-contact entry, and voice-first controls, Apollo Go makes urban mobility safer and more independent for visually impaired passengers—an often-overlooked group in traditional transport systems. These inclusive solutions help break down barriers to city life, promoting autonomy, social inclusion, and an improved quality of urban living that underpins resilient, accessible communities.

## 2.3 Future Work

Apollo Go is committed to continuously enhancing its accessibility features.

The platform will continue to collect feedback from visually impaired users to drive feature upgrades and improve its priority dispatch service for accessibility rides—helping to increase travel efficiency. Apollo Go also plans to expand the number of dedicated pick-up points for visually impaired passengers, based on user input, to better serve frequent travel scenarios and provide more convenient and inclusive mobility services.

Apollo Go is dedicated to meeting the diverse mobility needs of the larger 1.4 billion population in China. In addition to supporting visually impaired passengers, Apollo Go is also focused on the unique travel needs of women, youth, children with autism, seniors, and sanitation workers.

- Recognizing women's concerns about travel safety, Apollo Go has partnered with selected women community organizations to launch the 'Safe Travel' public welfare campaign. The initiative aims to establish supportive night-time stations that offer protection and reassurance for women during evening commutes.
- For teenagers, Apollo Go supports their interest in science and experiential learning by integrating autonomous driving technology into educational settings—offering real-life case studies and practical platforms for STEM (Science, Technology, Engineering, Mathematics) education.
- Understanding the needs of children with autism for sensory-friendly environments and social inclusion, Apollo Go uses autonomous vehicles to deliver a new kind of mobility experience—combining accessibility with advanced technology to foster both comfort and engagement.
- Apollo Go also addresses the needs of elderly passengers and sanitation workers. The company launched a special "Respect for Seniors Week," offering free rides to individuals over 60 years old. During the hot summer months, Apollo Go also established cooling stations to support frontline sanitation workers, providing a moment of relief amid challenging outdoor conditions.

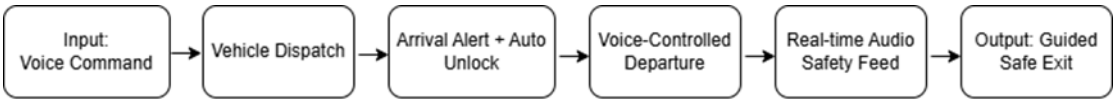
Apollo Go is committed to bringing warm, inclusive technology to everyone through human-centered innovation.

## 3 Use Case Requirements

Apollo Go's autonomous driving technology and sixth-generation autonomous vehicles ensure safe and reliable barrier-free travel. To achieve this, the team needs to fulfil below requirements:

- **REQ-01**: It is critical that autonomous driving systems possess a robust and scalable core to support widespread real-world deployment. Building on over a decade of research and operational expertise, Apollo Go has restructured its technology stack to center around advanced vision-based models. Apollo ADFM—the Autonomous Driving Foundation Model—is the first of its kind globally to enable Level 4 autonomous driving at scale, delivering the intelligence and adaptability required for complex urban scenarios.
- **REQ-02**: It is critical that safety frameworks in autonomous vehicles exceed traditional standards to ensure dependable operation under all conditions. The sixth-generation Apollo Go robotaxi introduces a pioneering 10-layer redundancy architecture, paired with a six-layer Multi-level Redundancy Control (MRC) system. These deeply integrated safety mechanisms reduce the probability of system failure to less than $10^{-8}$, approaching the aviation-level reliability of China's C919 commercial aircraft.
- **REQ-03**: It is critical that autonomous vehicles maintain full situational awareness in dynamic and adverse environments. The sixth-generation vehicle integrates 40 high-performance sensors across 7 categories, enabling five-layer, 360-degree environmental perception with a detection range of up to 440 meters. An adaptive cleaning system ensures continuous sensor functionality in rain, snow, and fog, empowering the vehicle to see clearly even when human visibility is compromised.
- **REQ-04**: It is critical that autonomous systems process massive volumes of data in real time to make precise, life-critical decisions. The Apollo Go robotaxi is equipped with dual AI multicore chips delivering over 1,200 trillion operations per second (TOPS). This high-performance computing platform functions as the vehicle's intelligent "brain," enabling faster, more accurate decisions than those made by human drivers.
- **REQ-05**: It is critical that autonomous vehicles provide intuitive and accessible user experiences to accommodate a wide range of passengers. The sixth-generation Apollo Go robotaxi integrates Baidu's ERNIE large language model to enable sophisticated, natural voice interaction. Passengers can adjust cabin conditions—including seat massage, temperature, windows, and music—through voice commands. The system also supports complex conversational Q&A, enhancing accessibility for visually impaired users and overall ride comfort for all.

## 4 Sequence Diagram




## 5 References

[1] Tencent Medical Dictionary. (2023, May 12). Current status of China's visually impaired population. WeChat Article "Tencent Medical Dictionary." https://mp.weixin.qq.com/s/sFFtKRIU49SQqTPzaBNm6A

[2] China Disabled Persons' Federation. (2022, August 3). Research on social inclusion of visually impaired individuals. WeChat Article "China Disabled Persons' Federation." https://mp.weixin.qq.com/s/-I8K3RWT4EU458Aaouvuuw

# Use Case 3: CargoWaves: The AI-Driven Inland Transport Platform Designed to Transform Logistics and Transport Industry

**Organization:** DP World

**Contact Person(s):**

Sabeen Muhammad – Sabeen.Muhammad@dpworld.com, +971 4 8898330, +971-55-2794335

Marwa Iqbal Mir Khoory – Marwa.Khoory@dpworld.com, +971-4-8896789

## 1 Use Case Summary Table

| Item | Details |
|---|---|
| Category | Intelligent transport |
| Problem Addressed | Previously, inland transport was plagued by inefficiencies such as manual truck booking, high dependency on brokers, limited real-time tracking, and excessive empty runs. These inefficiencies led to increased costs, delays, and environmental impact. By leveraging AI to optimize transport operations, Cargowaves eliminates fragmentation, reduces costs, and enhances sustainability, shaping the future of smart logistics. |
| Key Aspects of Solution | • Optimization of the booking process<br>• Tracking of the movement container from loading to delivery.<br>• Improved visibility of truck movements in the port for the terminal operations department.<br>• Enhanced milestone notification to all the relevant parties involved.<br>• Improved driver and terminal ops communication for support purposes.<br>• The key aspects of AI used in the solution were the Haulier recommendation model, Time slot Recommendation Models |
| Technology Keywords | Python, Rust, NextJS, PostgreSQL, XGBoost |
| Data Availability | Private |
| Standards | We adhere to ISO 27001 for ISMS in implementing our solutions & CRISP(Cross industry Standard Process)-ML (for AI model development)<br>Input – Text-based via web |
| Metadata (Type of Data) | Text (Relational Data) |

(continued)

| Item | Details |
|---|---|
| Model Training and Fine-Tuning | Both models are implemented using Python and Rust. SQL queries are used to fetch and match relevant data from the operational systems. The models are retrained and updated periodically to keep the recommendations relevant and accurate.<br>Model 1: Haulier Recommendation Model<br>This model uses XGBoost, a gradient boosting algorithm, to recommend the top 3 hauliers based on historical bookings and container specifications. The model pipeline includes:<br>• Data fetching and aggregation across multiple systems.<br>• Label encoding of categorical values (e.g., haulier names, container types).<br>• Handling of data imbalance using SMOTE (Synthetic Minority Over-sampling Technique).<br>• Model training using XGBoost, with iterative tuning until the F1 score exceeds 0.6.<br>If the accuracy is acceptable, feature importance is extracted to update attribute weights in the database. The model is retrained every 15 days to maintain accuracy.<br>Model 2: Time Slot Recommendation Model<br>This model uses a ratio-based scoring algorithm to recommend the best time slot for haulier execution. The process includes:<br>• Standardizing available slots by computing their proportion out of the total daily slots.<br>• Standardizing haulier execution data similarly.<br>• Computing probability scores by multiplying standardized slot values with execution ratios.<br>• Recommending the slot with the highest weighted score. This approach does not use a machine learning model but applies a mathematical policy-driven method using ratio-proportion and weighted averages to ensure fairness and slot efficiency. |
| Testbeds or Pilot Deployments | Pilot deployments & Live: dubaitrade.ae (test environment)<br>live environment[1] |

# 2 Use Case Description

## 2.1 Description

Cargowaves, an AI-driven inland transport platform, is transforming logistics by eliminating inefficiencies and enhancing cost-effectiveness for shippers, transporters, and cargo depots. Acting as a centralized digital marketplace, it streamlines truck booking, scheduling, and execution using AI-powered recommendations. This optimizes trucking operations, minimizes empty runs, and enhances supply chain visibility.

### Key Features & Benefits

- **AI-Powered Hauler Matching:** Recommends transport providers based on experience and price, improving cargo delivery efficiency.

- **End-to-End Digital Order Management:** Automates requests to final delivery, reducing manual processes.
- **Load Consolidation Optimization:** AI reduces empty haulage by 30%, maximizing truck utilization.
- **Real-Time Tracking:** Enhances cargo visibility, improving yard planning and traffic management.
- **AI-Based Terminal Slot Recommendations:** Ensures just-in-time communication, reducing wait times and enabling faster truck turnaround (up to 50%).
- **Automated Order Processing:** Cuts processing time by 50% by auto-drafting cargo orders, reducing redundancy and errors.
- **Digital Truck Booking:** Reduces booking time by up to 90% through automation.

### Technical explanation:

- The algorithm used was the XGBoost classifier, which is based on supervised learning
- The Input features are related attributes that impact the selection of the hauliers. The following features, container_iso_code, container_size, container_iso_type, and container_weight, have been identified as key features for matching a hauler with a Cargo owner
- The historical data is related to transactions done in the port by the hauliers for various Cargo Owners
- The SMOTE technique was used to address the issue of imbalanced data in question.
- We adhere to ISO27001, ISO20000 ,and Dubai Data law

### Impact on transport:

The initiative supports responsible and efficient resource use by integrating intelligent transport systems that reduce waste and optimize operations across logistics and mobility networks. By applying AI and data-driven decision-making, it ensures more efficient fuel use, reduced emissions, and better asset utilization, promoting sustainable practices in transport and infrastructure.

It also contributes to creating safer and more productive working environments through automation and intelligent monitoring. Enhanced visibility into transport operations and worker safety enables faster response to hazards and supports well-being in industrial and urban settings.

By embedding cutting-edge technologies—such as AI, IoT, and real-time analytics—into transportation systems, the initiative drives innovation across infrastructure and mobility sectors. It helps develop resilient, adaptable, and future-ready networks capable of handling growing urban and industrial demands.

Cargowaves enhances efficiency by automating workflows and leveraging AI-driven scheduling, reducing manual labor and errors by up to 80%. Faster transaction times– improved by 75%–boost economic productivity, enabling transporters to scale operations. The platform accelerates cargo throughput by 30%, lowers transport costs, and cuts idle truck hours by 50%, strengthening regional trade and business competitiveness.

AI-powered load matching and 100% paperless workflows modernize inland transport. By minimizing empty hauls by 30%, Cargowaves enhances industry resilience and optimizes infrastructure usage. Its predictive analytics help terminals manage traffic patterns, reduce idle time, and improve port efficiency, fostering a smarter, more sustainable logistics network.

Cargowaves minimizes empty hauls by 30%, significantly reducing fuel consumption and carbon emissions. The 100% digital workflow eliminates paper waste, lowering administrative overhead. Automated scheduling reduces congestion and idle time by 50%, leading to more sustainable resource consumption. The platform contributes to a 20% CO2 emissions reduction per shipment, driving eco-friendly logistics operations.

As a collaborative logistics marketplace, Cargowaves connects transporters and cargo owners, ensuring competitive pricing, enhanced visibility, and flexible cargo movement. Transporters benefit from reduced idle hours and increased revenue opportunities, while customers gain access to efficient and cost-effective transportation services, fostering a win-win ecosystem.

### Lessons learned from production deployment:

1. Due to compliance and data privacy requirements, the initial data shared was limited, making it challenging to build an accurate model.
2. During exploratory data analysis, we found that the dataset was highly imbalanced. To address this, we used techniques such as SMOTE and random oversampling to improve the performance of the boosting algorithm.
3. Considering the limited data that satisfied the criteria of the model, identifying the right accuracy metric was also challenging. Arriving at a reliable F1 score required multiple rounds of cross-validation.

## 2.2 Benefits of use case

Cargowaves enhances efficiency by automating workflows and leveraging AI-driven scheduling, reducing manual labor and errors by up to 80%. Faster transaction times– improved by 75%–boost economic productivity, enabling transporters to scale operations. The platform accelerates cargo throughput by 30%, lowers transport costs, and cuts idle truck hours by 50%, strengthening regional trade and business competitiveness.

AI-powered load matching and 100% paperless workflows modernize inland transport. By minimizing empty hauls by 30%, Cargowaves enhances industry resilience and optimizes infrastructure usage. Its predictive analytics help terminals manage traffic patterns, reduce idle time, and improve port efficiency, fostering a smarter, more sustainable logistics network. Cargowaves minimizes empty hauls by 30%, significantly reducing fuel consumption and carbon emissions. The 100% digital workflow eliminates paper waste, lowering administrative overhead. Automated scheduling reduces congestion and idle time by 50%, leading to more sustainable resource consumption. The platform contributes to a 20% CO2 emissions reduction per shipment, driving eco-friendly logistics operations.

As a collaborative logistics marketplace, Cargowaves connects transporters and cargo owners, ensuring competitive pricing, enhanced visibility, and flexible cargo movement. Transporters benefit from reduced idle hours and increased revenue opportunities, while customers gain access to efficient and cost-effective transportation services, fostering a win-win ecosystem.

## 2.3 Future work

AI-Driven Logistics & Digital Standards for CargoWaves

- Model Development: AI-Powered Predictive Logistics & Sustainability Optimization CargoWaves is enhancing efficiency, reducing emissions, and optimizing costs through AI-driven models.
- Predictive Demand Forecasting: Machine learning anticipates cargo fluctuations, enabling proactive truck/container allocation and real-time port slot scheduling, reducing congestion and idle time.
- Carbon Footprint Monitoring: AI-based emissions tracking provides real-time $CO_2$ calculations per shipment, offering actionable insights. AI recommends optimal routes, improving load consolidation and carbon KPIs.
- Safety & Traffic Monitoring: AI integrates with truck sensors and port safety systems to reduce accidents, optimize lane use, and enhance traffic flow. Smart safety scores assess transporters based on compliance, delivery efficiency, and driving behavior.
- Port AI Support Agents: AI-driven assistants provide predictable wait times, automated query resolution, and real-time issue handling to enhance the driver-terminal experience.

Standards Development: Regulatory & Industry Guidelines for Digital Logistics

To accelerate the transport industry's shift towards digitalization, CargoWaves will focus on developing standardized frameworks that promote seamless technology adoption, AI-driven communication, and data interoperability across stakeholders.

- **E-Payments & Digital Contracts**: Encouraging cashless transactions for secure and seamless trade.
- **AI-Powered Communication Protocols**: Enabling real-time driver-port interactions for better slot management and cargo tracking.
- **AI-Based Predictive Scheduling**: Automating truck deployment based on demand patterns to optimize transport operations.

By focusing on AI-powered optimization and standardized digital logistics frameworks, CargoWaves can drive sustainable, efficient, and globally scalable supply chain transformation.

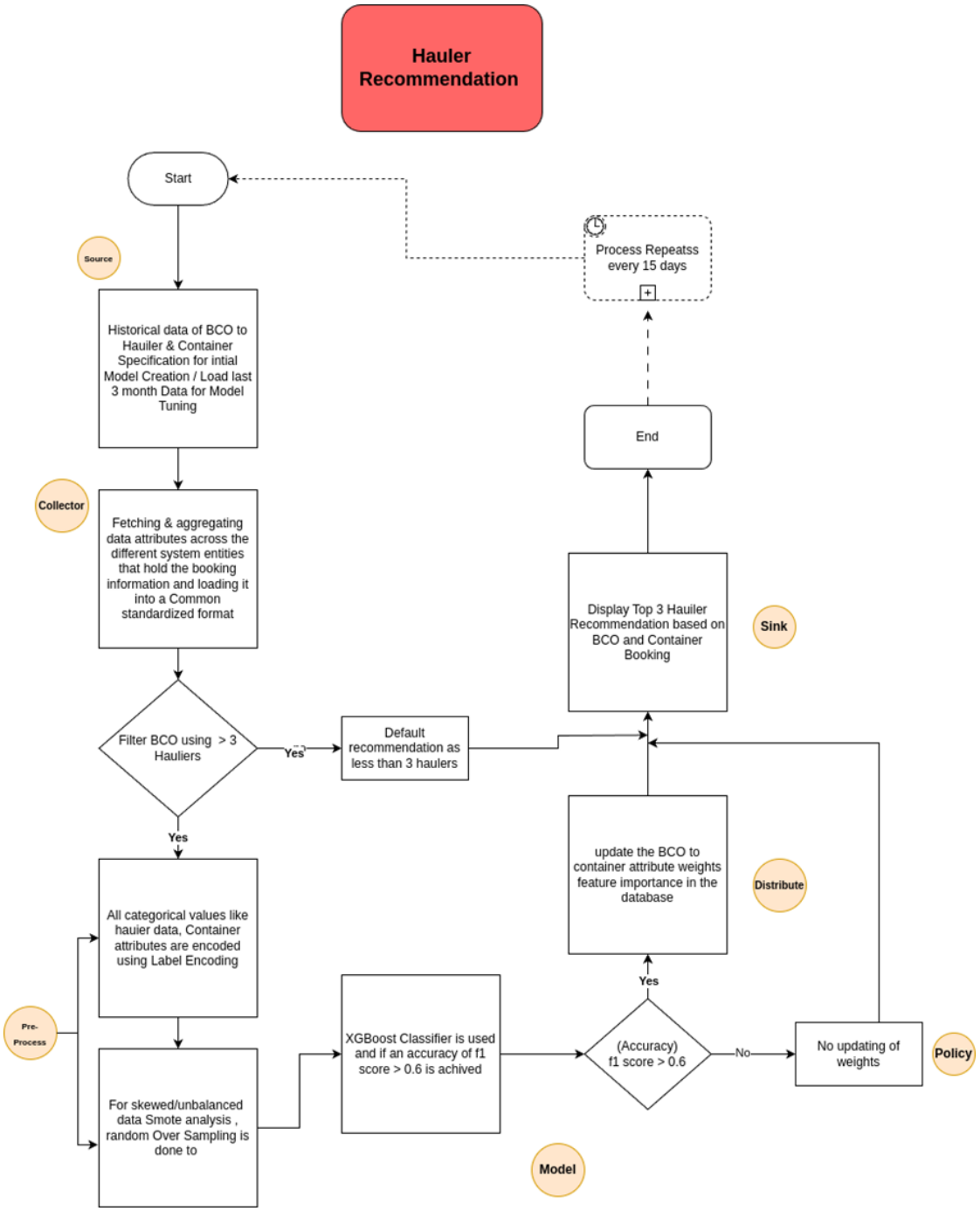
### Deployment & Usage

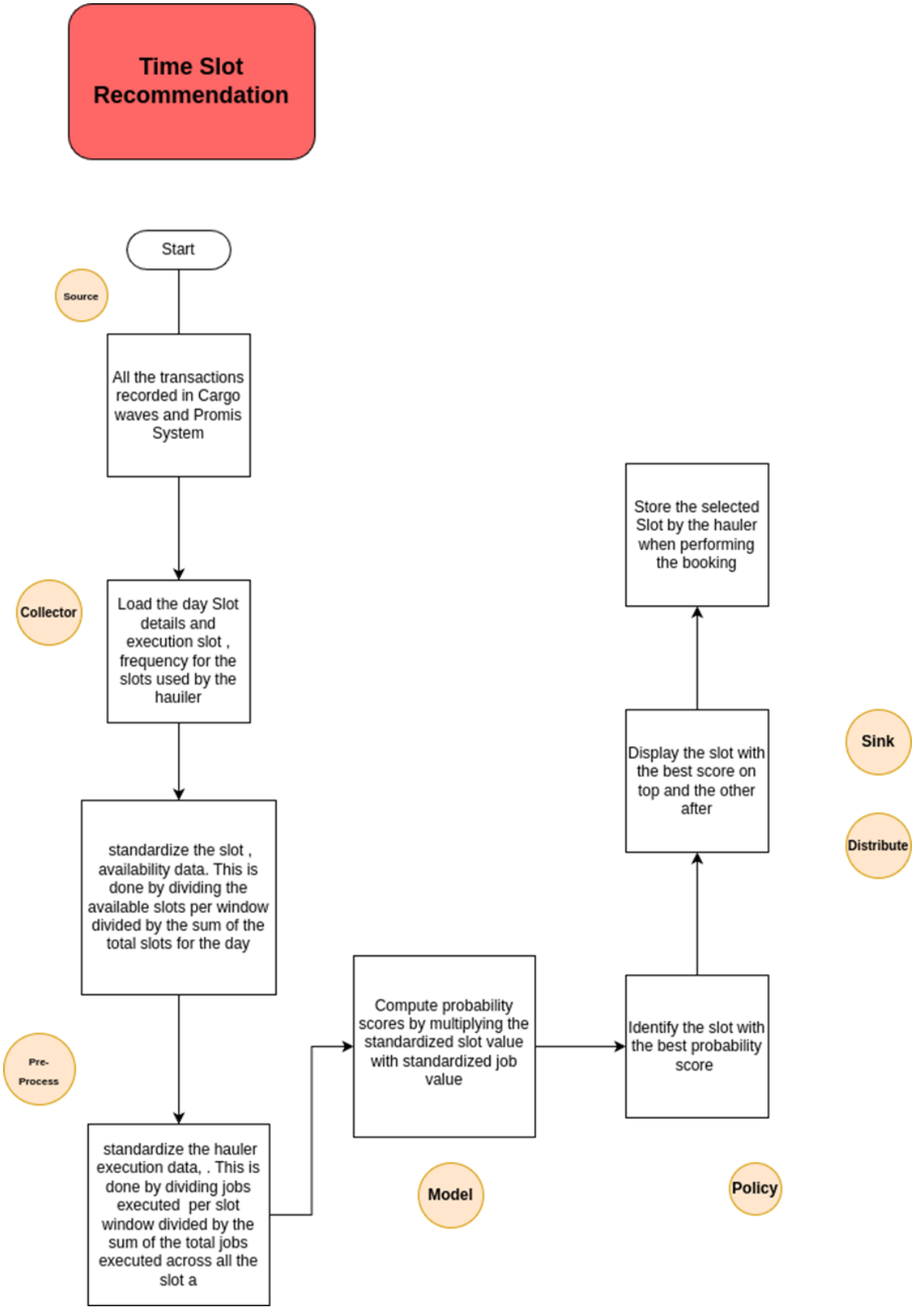
**From the Recommendation systems:** 15% of the recommendations have been used by the Beneficiary Cargo Owners for completing their booking.

# 3 Use Case Requirements

- **REQ-1**: **It is critical** to display a recommended hauler best suited to fulfill the current booking, based on historical interactions between the BCO and the hauler community. The recommendation system must prioritize boosting the most relevant hauler. The trained data must be stored in a relational database and accessed via SQL queries, with the interface to the front end implemented through REST APIs. **It is mandatory** that the model is retrained every 15 days and that all selections made from the recommendations are recorded in the system.
- **REQ-2**: **It is critical** to display a recommended timeslot based on the current port availability and historical slot usage by the BCO. The recommendations must be computed in real time using probabilistic techniques and weighted averages.

# 4 Sequence Diagram




Time Slot Recommendation
Start
Source
All the transactions recorded in Cargo waves and Promis System
Collector
Load the day Slot details and execution slot , frequency for the slots used by the hauiler
standardize the slot , availability data. This is done by dividing the available slots per window divided by the sum of the total slots for the day
Pre-Process
standardize the hauler execution data, . This is done by dividing jobs executed per slot window divided by the sum of the total jobs executed across all the slot a
Compute probability scores by multiplying the standardized slot value with standardized job value
Model
Identify the slot with the best probability score
Policy
Display the slot with the best score on top and the other after
Sink
Distribute
Store the selected Slot by the hauler when performing the booking

## 5 References

[1] Dubai Trade, "Dubai Trade Portal," Dubai Trade, [Online]. Available: https://www.dubaitrade.ae/en/.

[2] Nagpal, Gaurav, et al. "Use of data analytics to increase the efficiency of last mile logistics for ecommerce deliveries." Handbook of research on engineering, business, and healthcare applications of data science and analytics. IGI Global Scientific Publishing, 2021. 167-180.

[3] Wang, Xing, et al. "An XGBoost-enhanced fast constructive algorithm for food delivery route planning problem." Computers & Industrial Engineering 152 (2021): 107029.

[4] Arifianti, Fidela Putri, and Abu Salam. "XGBoost and Random Forest Optimization using SMOTE to Classify Air Quality." Advance Sustainable Science, Engineering and Technology 6.1 (2024): 1-8.

[5] Imani, Mehdi, Ali Beikmohammadi, and Hamid Reza Arabnia. "Comprehensive analysis of random Forest and XGBoost performance with SMOTE, ADASYN, and GNUS under varying imbalance levels." Technologies 13.3 (2025): 88.

# Use Case 4: Low-Cost Decision Making for Group of Vehicles for Safe Mobility in V2X

**Organization:** Tata Consultancy Services Limited

**Country**: India

**Contact Person(s):**

Primary contact: Garima Mishra, garima.mishra2@tcs.com, 9811964569

Secondary contact: Sai Bhavadeesh Yarlagadda, saibhavadeesh.yarlagadda@tcs.com, 8106263461,

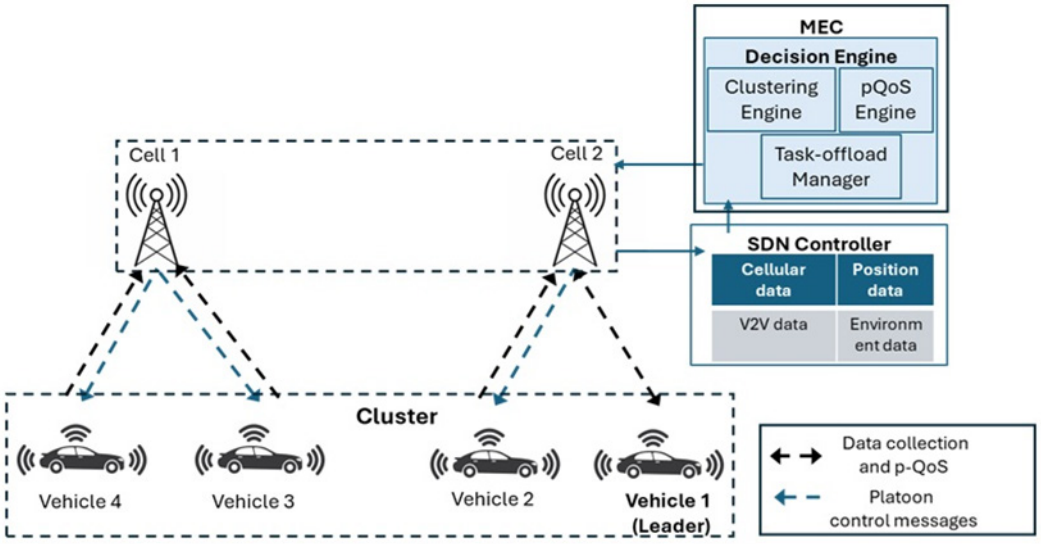
Hemant Kumar Rath, hemant.rath@tcs.com, 9035017741

## 1 Use Case Summary Table

| Item | Details |
|---|---|
| Category | Low-cost and Safe Transport, V2X, 5G/6G |
| Problem Addressed | 1. Change in Quality of Service (QoS) drastically within a short period due to high mobility and fading of the wireless channel,<br>2. Pre-programmed ML algorithms,<br>3. High cost of on-board data computation |
| Key Aspects of Solution | Enhance the safety of vehicles, reduce the cost of expensive on-board data computation and analytics, and predictive QoS for vehicle platooning in multi-environment scenarios. |
| Technology Keywords | C-V2X, Predictive-QoS, multi-environment, Explainable AI (XAI), Vehicle platooning |
| Data Availability | Public dataset: 1. Berlin-V2X Dataset[1]<br>2. TiHAN-V2X Dataset[2] |
| Metadata (Type of Data) | Time-series data |
| Model Training and Fine-Tuning | Data pre-processing (Data cleaning, Feature reduction and selection), Lag-variable optimization, Trained on LightGBM Regressor for real-time deployment, SHAP analysis for Explainable AI (XAI). |
| Testbeds or Pilot Deployments | Virtual simulations in controlled environments |
| Code repositories | Fraunhofer HHI 2024 [9]<br>Takalani95 2024 [10] |

# 2 Use Case Description

## 2.1 Description




In the emerging world of intelligent transportation systems, Vehicle-to-Everything (V2X) communication is evolving to provide safer and more efficient road networks. Some of the promising applications of V2X for wireless control use cases are Platooning, Cooperative Transport of Goods, Cooperation between Automated Guided Vehicles (AGVs) and Stationary Robots, etc. [1]. These applications require frequent exchange of data between vehicle sensors and wireless networks to fulfill Quality of Service (QoS) requirements for safe and coordinated movement.

We consider a platooning of vehicles scenario, as shown in Fig. 1, in which several vehicles are traveling in a group, along a highway, and connected to base stations. The vehicles within the group trail a single leading vehicle positioned ahead of them. Data from these vehicles, along with network information from their associated base stations, is gathered at a Software Defined Network (SDN) controller. The data includes wireless network information such as base station transmit power, bandwidth, wireless link quality in terms of RSSI, RSRP, RSRQ, and SNR, etc, and metadata such as vehicle-to-vehicle (V2V) data: vehicle location information, sensor data: camera data and environment data: weather condition, traffic jam factor, etc. This extensive data is used for making safety decisions at the Multi-access Edge Computing (MEC) to minimize the risk of collisions or congestion.

### Problem:

The current platooning solutions often rely on expensive and complex decision-making algorithms for coordinated movement. Hence, QoS estimation in the form of reliable wireless transmission is important for wireless control use cases. QoS estimation in non-stationary radio environments can help to avoid any unexpected service interruptions by allowing the network to take proactive actions. It helps vehicles to take actions such as task offloading of on-device computation and decision making. Examples of task offloading include gaming, object detection, and infotainment services. Proactive actions are taken based on inferences,

contributing to crash prevention, enhancing vehicle safety, and uninterrupted communication. However, proactive decision-making in such systems presents several challenges:

1. Mixed Traffic Conditions & Multi-Environment Scenarios: Roads are unpredictable with human-driven vehicles, pedestrians, and unexpected obstacles influencing the platoon's stability. Network link conditions are highly impacted by multi-environment scenarios such as highways, avenues, residential areas, and park, etc.
2. Pre-programmed ML Models: Most of the applications today use pre-programmed ML models, which are designed for specific scenarios, limiting their adaptability to diverse mobility conditions and varying network environments.
3. Computational Costs: The cost of processing vehicular sensor data on board varies widely depending on factors such as model complexity, the number of sensors used, processing power required, and hardware capabilities, etc. While basic processing can cost only a few cents per gigabyte, advanced computing needs, such as those for autonomous driving, can reach tens of dollars per gigabyte, posing a significant challenge for real-time QoS predictions.

### Need for stationarity check and predictive QoS for vehicle safety:

Our use case focuses on developing a cost-effective decision-making system, ensuring safe mobility for vehicle platooning in a V2X communication network. It leverages real-time data exchange among vehicles, infrastructure, and networks. By utilizing the extensive data available, we enable proactive decision-making that satisfies End-to-End (E2E) QoS requirements. Proactive decision-making is based on a time stationarity check of multi-environment data and predictive QoS (pQoS), e.g., latency, packet loss ratio, or throughput [2].

The time stationarity is a process that checks for a change in the unconditional joint probability distribution of wireless network data with time. The pQoS is a metric that enables forecasting/predicting wireless network quality well in advance.

- The stationarity check information is used for the formation of the optimum platoon, ensuring safe distances between vehicles, and improving overall traffic efficiency.
- The E2E pQoS predicts the E2E quality of wireless connectivity of the communication link. The E2E pQoS takes care of task offloading decisions. Without pQoS, any offloading strategy would eventually result in service disruption.

By utilizing lightweight algorithms and edge computing, the system ensures affordability while improving road safety [3], cost-efficiency, and autonomous coordination in connected transportation environments.

**Objective:** To develop low-cost and safe mobility, End-to-End (E2E) pipelines for the reliable prediction of QoS parameters and task offloading for vehicle platooning in multi-environment scenarios by leveraging data from both mobile networks and vehicle sensors.

### Use Case Status: The use case is part of a larger research project.

### Detailed solution proposal

We address the challenge of unavailable or computationally expensive network data for QoS prediction by enhancing our Machine Learning (ML) model with prior information from a leading vehicle [4].

By analyzing correlations among vehicles traveling the same path, we observe that they experience similar large-scale fading effects, leading to strong correlations with the leading vehicle's measurements. This raises an important question: to what extent can past user experiences help predict the propagation environment? To explore this, we evaluate our ML model for QoS prediction not only with data collected from a single leading vehicle but also by incorporating insights from prior users. Our findings indicate that by leveraging data from the leading vehicle, QoS prediction accuracy improves significantly for following vehicles on the same route.

Furthermore, we employ Explainable AI (XAI) to demonstrate that ML models can learn wireless network conditions without pre-programming [5]. Our dataset is collected from a real-world network, covering platooning vehicular scenarios for multi-radio environments.

This work lays the foundation for reinforcing the following benefits:

1. Platooning promises higher fuel efficiency due to reduced aerodynamic drag and increased overall traffic efficiency because of smoother velocity control.
2. Clustering more vehicles in a platoon optimizes costly data collection and reduces the need for extensive data for pQoS evaluation
3. pQoS-based decision for full/partial task offloading of expensive computing tasks

## System Architecture:

Our proposed solution incorporates two E2E pipelines, along with several key modules, designed to optimize the decision-making process for safe and efficient vehicle mobility in a V2X network.

## Pipelines:

1. **Master Pipeline:** The master pipeline is responsible for data collection, cluster formation, and QoS prediction. The Software-Defined Network (SDN) Controller at the base station aggregates network data, including GPS information from the vehicles, and shares it with the clustering engine and the pQoS engine as shown in Fig. 1. First, the clustering engine performs a stationarity test on the data, forms a vehicle cluster, and elects a leader vehicle. The cluster data, including the number of vehicles and the leader vehicle's identity, is then sent to the respective vehicles. Later, the pQoS engine predicts the network's QoS and relays this information to the leader vehicle, which subsequently uses it to make task offloading decisions.
2. **Invoked Pipeline:** The invoked pipeline focuses on task offloading. The leader vehicle evaluates whether the predicted QoS is adequate to offload tasks to a task offloading manager co-located with a nearby Multi-Access Edge Computing (MEC). If the QoS meets the requirements, the leader offloads its computational tasks to the task offloading manager. These computationally intensive tasks are processed, and the task offloading manager generates a decision. This decision is sent back to the leader, who then communicates the outcome to the other vehicles in the cluster.

## Modules:

1. **Data Collector:** The SDN Controller at the base station functions as the Data Collector. It aggregates information from the vehicles, including GPS data, speed, and other relevant parameters, as well as network-related data.

2. **Decision Engine:** It contains three modules, such as clustering engine, pQoS engine, and task offloading manager.
    - **Clustering Engine:** The clustering engine is tasked with forming dynamic clusters, electing a leader vehicle, and ensuring safe distances between vehicles through stationarity tests. Once the cluster is formed and the leader is elected, the engine updates the network with this information. Since this process is ongoing, the clusters are dynamic and adapt to changing network conditions.
    - **pQoS Engine:** The pQoS Engine plays a critical role in predicting the QoS of the network, which is essential for making informed task offloading decisions. It employs ML models to forecast network behaviour and performance in the future. The offloading decision is based on an End-to-End (E2E) prediction of the vehicle-to-network link. Predictive QoS not only enables the feasibility assessment of task offloading but also ensures a seamless transition back to in-vehicle resources if the network conditions degrade.
    - **Task Offloading Manager:** Autonomous vehicles require substantial computational power for tasks like critical-safety applications. Due to space, cost, and energy constraints, partial or full offloading of these tasks to a cloud-based MEC environment is a viable solution. The task offloading manager is responsible for executing these tasks in the cloud and making decisions regarding task offloading. The task offloading manager sends the response back to the leader vehicle, which then communicates the decision to the other vehicles in the cluster, ensuring efficient resource management across the system.

## Impact on Intelligent Transport

The integration of AI into transport systems is transforming the future of mobility by directly enhancing road safety, operational efficiency, and user experience. A primary impact is the reduction of traffic-related deaths and injuries through intelligent systems that anticipate risks, improve driver behavior, and enable faster emergency response.

AI-powered tools improve the accuracy of crash data collection and analysis, offering insights that help prevent accidents and guide better infrastructure planning. These systems can support dynamic traffic management and enable predictive interventions that reduce congestion and increase overall transport system efficiency.

In the realm of intelligent transport, autonomous and semi-autonomous features such as lane-keeping assistance, emergency braking, and automatic parking are already making vehicles safer and easier to operate. These technologies also reduce driver workload and error, key factors in many road incidents.

Moreover, intelligent transport benefits from AI-enhanced traffic coordination, enabling real-time adjustments based on road conditions, vehicle density, and weather. This results in more responsive and resilient urban mobility networks.

Emerging mobility solutions—such as connected vehicles, ride-sharing platforms, and integrated infotainment and communication systems—further contribute to smarter transportation ecosystems. These innovations help streamline vehicle coordination, improve route efficiency, and create safer, more enjoyable travel experiences.

Ultimately, the adoption of AI in transport is not only about improving individual vehicle performance but about building a connected, adaptive, and sustainable transport infrastructure that enhances safety, accessibility, and urban quality of life on a system-wide level.

## 2.2 Benefits of use case

The integration of AI into transport systems is transforming the future of mobility by directly enhancing road safety, operational efficiency, and user experience. A primary impact is the reduction of traffic-related deaths and injuries through intelligent systems that anticipate risks, improve driver behavior, and enable faster emergency response.

AI-powered tools improve the accuracy of crash data collection and analysis, offering insights that help prevent accidents and guide better infrastructure planning. These systems can support dynamic traffic management and enable predictive interventions that reduce congestion and increase overall transport system efficiency.

In the realm of intelligent transport, autonomous and semi-autonomous features such as lane-keeping assistance, emergency braking, and automatic parking are already making vehicles safer and easier to operate. These technologies also reduce driver workload and error, key factors in many road incidents.

Moreover, intelligent transport benefits from AI-enhanced traffic coordination, enabling real-time adjustments based on road conditions, vehicle density, and weather. This results in more responsive and resilient urban mobility networks.

Emerging mobility solutions—such as connected vehicles, ride-sharing platforms, and integrated infotainment and communication systems—further contribute to smarter transportation ecosystems. These innovations help streamline vehicle coordination, improve route efficiency, and create safer, more enjoyable travel experiences.

Ultimately, the adoption of AI in transport is not only about improving individual vehicle performance but about building a connected, adaptive, and sustainable transport infrastructure that enhances safety, accessibility, and urban quality of life on a system-wide level.

## 2.3 Future work

1. **Data collection:** The current dataset lacks quality and quantity. Hence, in the future, for improved performance analysis, we need to have a dataset with a larger number of vehicles in different platooning scenarios.

2. Proof of concept development

- Solving AI/ML Domain Adaptation Through Transfer Learning – Domain adaptation is a key challenge in AI/ML, especially when models trained in one domain need to generalize effectively to a different but related domain. Transfer learning provides a powerful approach to address this issue by leveraging knowledge from a source domain to improve performance in a target domain with limited labelled data.
- Making AI-driven systems more cost-effective and scalable – Explainable AI (XAI) can help reduce computational costs in machine learning applications, including QoS prediction and vehicular communications, by enhancing model efficiency, interpretability, and optimization. With these advancements, the research can transition from theoretical modelling to practical, real-world applications, enhancing the security of V2X systems.

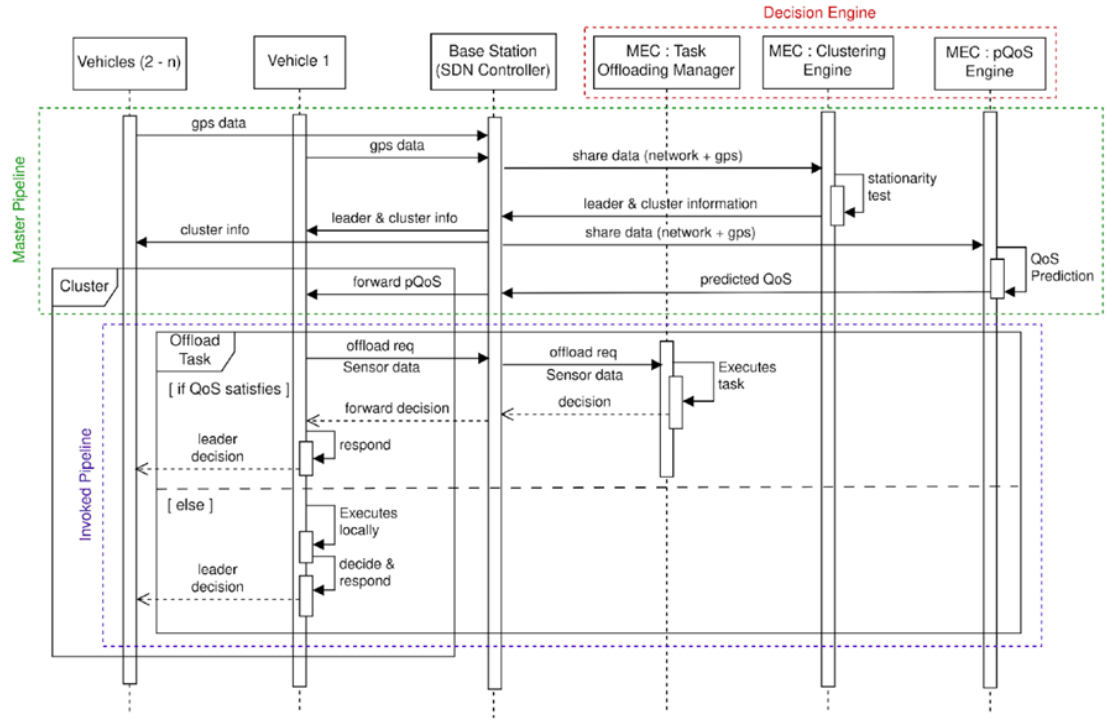
# 3 Use Case Requirements

- **REQ-01**: It is critical to enhance Quality of Service (QoS) by defining the minimum communication performance indicators, such as sub-10 millisecond latency and 99.999%

reliability[6], that 5G/6G networks and Multi-access Edge Computing (MEC) must meet to support real-time decision-making in vehicular environments.

- **REQ-02**: It is mandatory to clearly specify the type of cellular connectivity required, distinguishing between technologies like NR-V2X PC5 sidelink and Uu uplink, and to define associated bandwidth requirements and uplink reliability thresholds to ensure consistent communication performance.
- **REQ-03**: It is critical to differentiate between static data collection (for offline model training) and real-time telemetry (for online inference), while also detailing spatial-temporal data coverage expectations and data modalities including GPS, LiDAR, and RSSI.
- **REQ-04**: It is mandatory to define specific technical capabilities for the simulation platform, including support for multi-radio vehicular mobility, stochastic channel modeling, and integration with SDN/MEC emulators such as Mininet-WiFi, OMNeT++, or NS-3[6], to accurately emulate real-world transport scenarios.
- **REQ-05**: It is critical to separate and articulate the requirements for dynamic machine learning models, including support for online learning, explainability (via eXplainable AI/XAI), model quantization for efficient edge deployment, and real-time adaptability to domain shifts using techniques like transfer learning or continual learning.

## 4 Sequence Diagram




The workflow of the use case, as shown in Fig. 1, follows the sequence of tasks as shown in the above sequence diagram, Fig. 2. Initially, the SDN module inside the base station aggregates network data, including GPS information from the vehicles, and shares it with the clustering engine and the pQoS engine, which are hosted at the MEC. The clustering engine performs a stationarity test on the data, forms a vehicle cluster, and elects a leader vehicle (viz. Vehicle 1). The cluster data, including the number of vehicles and the leader vehicle's identity, is then sent to the respective vehicles. And the pQoS engine predicts the network's QoS and relays this information to the leader.

If the QoS received by the leader meets the threshold, then the leader sends an offloading request to the base station. The base station then communicates with the task offloading manager, located near the leader vehicle. The task offloading manager executes these tasks in the cloud and makes decisions regarding task offloading. The task offloading manager sends the response back to the leader vehicle, which then communicates the decision to the other vehicles in the cluster, ensuring efficient resource management across the system. If the QoS is not met, then the leader vehicle executes the task locally and makes decisions, which it then communicates the decision to other vehicles in the cluster.

This process ultimately facilitates low-cost decision-making and reduces the load on the on-device resources.

## 5 References


[1] Hernangomez, Rodrigo, Philipp Geuer, Alexandros Palaios, Daniel Schäufele, Cara Watermann, Khawla Taleb-Bouhemadi, Mohammad Parvini, Anton Krause, Sanket Partani, Christian Vielhaus, Martin Kasparick, Daniel F. Külzer, Friedrich Burmeister, Frank H. P. Fitzek, Hans D. Schotten, Gerhard Fettweis, and Slawomir Stanczak. "Berlin V2X." *IEEE Dataport*, December 8, 2022. https://doi.org/10.21227/8cj7-q373.

[2] Annu, Phaneendra Pydimarri, Pachamuthu Rajalakshmi, Sunnam Venkata Srikanth, and Munjagala Prasad. "TiHAN-V2X: A Comprehensive Dataset for Dynamic C-V2X Communication within the Indian Context." *IEEE Dataport*, October 8, 2024. https://doi.org/10.21227/f2kd-9g03.

[3] Kulzer, Daniel F., et al. "AI4Mobile: Use Cases and Challenges of AI-based QoS Prediction for High-Mobility Scenarios." April 2021. https://doi.org/10.1109/vtc2021-spring51267.2021.9449059.

[4] Palaios, Alexandros, et al. "Machine Learning for QoS Prediction in Vehicular Communication: Challenges and Solution Approaches." *IEEE Access* 11 (January 2023): 92459–92477. https://doi.org/10.1109/access.2023.3303528.

[5] International Telecommunication Union (ITU). "AI for Road Safety Global Initiative." 2025. https://www.itu.int/en/ITU-T/ITS/AIRoadSafety/Pages/default.aspx.

[6] Ain, N. U., Rodrigo Hernangómez, Alexandros Palaios, Martin Kasparick, and Slawomir Stańczak. "QoS Prediction in Radio Vehicular Environments via Prior User Information." *arXiv*, February 2024. https://doi.org/10.48550/arxiv.2402.17689.

[7] Adadi, Amina, and Mohammed Berrada. "Peeking Inside the Black-Box: A Survey on Explainable Artificial Intelligence (XAI)." *IEEE Access* 6 (October 2018): 52138–52160. https://doi.org/10.1109/access.2018.2870052.

[8] Annu, and P. Rajalakshmi. "Towards 6G V2X Sidelink: Survey of Resource Allocation—Mathematical Formulations, Challenges, and Proposed Solutions." *IEEE Open Journal of Vehicular Technology* 5 (2024): 344–383. https://doi.org/10.1109/OJVT.2024.3368240.

[9] Fraunhofer Heinrich Hertz Institute. *BerlinV2X*. GitHub. Accessed June 19, 2025. https://github.com/fraunhoferhhi/BerlinV2X.

[10] Takalani95. *QoS Prediction Challenge by ITU AI/ML in 5G Challenge*. GitHub. Accessed June 19, 2025. https://github.com/takalani95/QoS-Prediction-Challenge-by-ITU-AI-ML-in-5G-Challenge.

# Use Case 5: Intelligent In-Vehicle Scenario Orchestration Agent System

**Country**: China

**Organization**: Chongqing Changan Automobile Co. Ltd.

**Contact Person:**

Yonggang Luo(luoyg3@changan.com.cn)

## 1 Use Case Summary Table

| Item | Details |
|---|---|
| Category | Intelligent transport |
| Problem Addressed | Traditional vehicle control configuration processes are cumbersome, fragmented, and heavily reliant on professional knowledge, requiring users to invest significant time and effort to complete tasks. |
| Key Aspects of Solution | The system leverages large language models to accurately recognize and understand user intentions. It automatically generates logical chains containing "preconditions", "triggers", and "executions", significantly improving scenario orchestration efficiency and scalability. |
| Technology Keywords | Large Language Models, Service Orchestration. |
| Data Availability | Private |
| Metadata (Type of Data) | Text<br>Input data: Chinese user intents<br>Output data: Chinese and English car services |
| Model Training and Fine-Tuning | SFT With approximately 20 million tokens of automotive domain data and around 400,000 instruction tuning data. |
| Testbeds or Pilot Deployments | TikTok Video [1] |
| Code repositories | None |

## 2 Use Case Description

### 2.1 Description

Intended use: The system allows users to dynamically create and manage complex in-vehicle scenario modes, such as personalized climate control and entertainment settings, using natural

language input via voice. This significantly reduces task execution time from minutes to seconds, streamlining the overall driving experience.

Problem to solve: Conventional vehicle control configuration is time-consuming and unintuitive, often requiring multiple manual inputs across isolated systems. These processes typically take several minutes and demand technical knowledge, posing usability challenges for non-expert users.

Limitations of existing solutions: Existing solutions rely heavily on static presets or manual configuration interfaces, offering limited adaptability to user preferences or dynamic environments. They lack the capability to orchestrate functions across subsystems in real-time, and often cannot integrate seamlessly with legacy vehicle platforms or third-party services.

Benefits and drawbacks of the AI-based approach: The system leverages a large language model trained on approximately 20 million automotive tokens and 400,000 task-specific instructions to parse user input and autonomously generate scenario logic chains comprising preconditions (e.g., vehicle speed < 30 km/h), triggers (e.g., voice command detected), and executions (e.g., activate low-speed cruise mode). This enables context-aware, modular orchestration of complex vehicular functions.

Benefits include:

1. High-performance natural language understanding for domain-specific queries.
2. Dynamic invocation of cross-system functions (e.g., ADAS, infotainment, HVAC).
3. Integration with SparkLink to enable low-latency, high-reliability short-range communication.

Limitations include:

1. Dependence on training data quality poses risks for edge cases and uncommon driving scenarios.
2. Ongoing challenges in model interpretability, error recovery, and adaptation to real-time edge environments.

### Impact on Intelligent Transport

This system supports global efforts toward sustainable development by promoting more efficient, eco-friendly, and innovative transportation solutions. By enabling flexible scenario orchestration and automated vehicle control, it significantly reduces energy waste, enhances vehicle usage efficiency, and promotes environmentally responsible travel behavior.

The integration of advanced artificial intelligence models with digital twin technology allows for real-time optimization of resources and decision-making. This leads to a safer, more comfortable, and sustainable driving experience. Modeling entire vehicle services in the cloud for training and testing AI systems strengthens the foundation of smart transportation networks.

These innovations contribute to building sustainable and resilient urban mobility, supporting climate-conscious strategies, and fostering technological advancement. Ultimately, this approach enhances the quality of life in cities and promotes the broader vision of intelligent, green, and inclusive transportation systems.

Partner name: Ali Cloud Computing Co. Ltd. (Collaborating on LLMs technologies), IEEE SA (Collaborating on the standard for MLLMs platforms)

## 2.2 Benefits of use case

Benefits include:

1. High-performance natural language understanding for domain-specific queries.
2. Dynamic invocation of cross-system functions (e.g., ADAS, infotainment, HVAC).
3. Integration with SparkLink to enable low-latency, high-reliability short-range communication.

This system supports global efforts toward sustainable development by promoting more efficient, eco-friendly, and innovative transportation solutions. By enabling flexible scenario orchestration and automated vehicle control, it significantly reduces energy waste, enhances vehicle usage efficiency, and promotes environmentally responsible travel behavior.

The integration of advanced artificial intelligence models with digital twin technology allows for real-time optimization of resources and decision-making. This leads to a safer, more comfortable, and sustainable driving experience. Modeling entire vehicle services in the cloud for training and testing AI systems strengthens the foundation of smart transportation networks.

These innovations contribute to building sustainable and resilient urban mobility, supporting climate-conscious strategies, and fostering technological advancement. Ultimately, this approach enhances the quality of life in cities and promotes the broader vision of intelligent, green, and inclusive transportation systems.

## 2.3 Future work

In the next 6–12 months, efforts will center on foundational integration and prototyping of the system's key components:

**Multimodal Sensor & Environment Integration:** Expand the system's perception by incorporating a wider array of vehicle sensor inputs and contextual environmental information. These additions will be integrated into the digital twin orchestration platform to simulate rich, realistic scenarios for development and testing. Early prototypes will focus on fusing these diverse data sources in real time, improving situational awareness and enabling the digital twin to mirror real-world conditions for robust scenario planning.

**LLM–Agent Interface Development:** Establish a communication framework that enables LLMs to interface with autonomous driving agents. This involves designing a clear interface protocol through which the LLM can receive processed perceptual data and issue driving directives or queries. An initial version of an intent interpretation module will be implemented to translate the LLM's natural-language or abstract directives into structured, machine-interpretable commands for vehicle control. Leveraging insights from recent studies showing that LLMs can transform high-level instructions into precise, executable action sequences for robots, the interface will be designed to support such translation of intent into action for autonomous vehicles. Basic action delegation mechanisms will be prototyped, wherein the autonomous agent accepts LLM commands and carries them out via low-level controllers. A rudimentary feedback loop will

also be established: the agent will return outcome signals to the LLM or a supervising module, enabling iterative refinement of commands even at this early stage.

**Pilot Federated Learning & Privacy Baseline:** Initiate an FL pilot using the digital twin platform to connect multiple virtual vehicles for collaborative model training. In the short term, the focus is on proving the feasibility and benefits of distributed learning across agents without centralizing raw sensor data. Key performance benchmarks will be measured. For example, using FL, only model parameters or gradients are shared rather than raw images or point clouds, dramatically reducing communication loads. The initial goal is to achieve near real-time aggregation intervals and ensure the global model converges to within an acceptable range of a centrally trained model's accuracy while significantly minimizing the data sent over the network. In parallel, foundational privacy-preserving techniques will be applied. A modest differential privacy guarantee may be introduced to gauge its impact on model performance. By the end of this phase, the project should demonstrate a working FL prototype where each vehicle's sensor data remains local, basic privacy measures are in place, and the system can learn a simple shared model with measurable efficiency and privacy benefits.

In the 1–2 year timeframe, the project will build on the prototypes to scale up the capability, robustness, and sophistication of the intelligent mobility ecosystem:

**Enhanced Sensor Fusion & Environmental Context:** Integrate additional data modalities and external data feeds into both the vehicle and its digital twin. Mid-term goals include incorporating V2X inputs from smart infrastructure directly into the vehicle's decision-making loop. The digital twin orchestration platform will be extended to model complex urban environments, including dynamic traffic conditions and weather events, allowing validation of the system under a wide range of scenarios. By combining on-board sensor fusion with V2X communication, the autonomous agent will gain the ability to see beyond line-of-sight, improving safety and efficiency. The digital twin will serve as a testbed for these features, ensuring that the fusion of vehicle sensors with external inputs is calibrated and reliable before deployment.

**Robust LLM–Agent Collaboration:** With a basic LLM-agent interface in place, the mid-term work will refine and formalize this integration. A unified protocol will be developed in collaboration with industry standards. The intent interpretation module will be made more sophisticated, possibly using semantic parsing or intermediate planning languages to accurately capture the LLM's intent. For instance, if the LLM suggests a complex maneuver, the system will break this down into actionable sub-tasks that the agent can execute. Improved action delegation strategies will be implemented: the autonomous agent will intelligently allocate tasks between its traditional control stack and the high-level LLM-driven planner. Safety and explainability will be prioritized. Concurrently, feedback loops between the agent and LLM will be strengthened. The vehicle agent will continuously feed back state updates to the LLM, enabling the LLM to adjust its plans in real time. This interactive loop aligns with techniques like grounded decoding in robotics, wherein LLM plans are dynamically conditioned on the physical environment's state. We anticipate that by the end of two years, the LLM-assisted planning system will be capable of handling complex, multi-step driving tasks in simulation with high success rates, thanks to a well-defined interface and continuous agent–LLM communication.

**Federated Learning at Scale with Privacy Guarantees:** The mid-term milestone for federated learning is to scale up the number of participating vehicles. Effort will be devoted to optimizing the convergence rate of the federated model. Success will be measured by the global model

reaching a high accuracy within a limited number of communication rounds, and maintaining stability even as more vehicles join the training. Concretely, the aim is to approach centralized training accuracy to within a few percent, with minimal increase in training time. Techniques like personalized federated learning will be explored to accommodate differences between vehicles. In terms of communication efficiency, the system will incorporate model update compression and update frequency reduction, targeting an order-of-magnitude reduction in required bandwidth versus naive approaches. Privacy-preservation will be substantially reinforced: DP will be tuned to provide quantifiable privacy guarantees. The chosen DP noise levels will be calibrated so that the model's utility remains high while ensuring that individual data points from any single vehicle cannot be inferred from the shared updates. In addition, more advanced privacy-preserving techniques will be prototyped. The mid-term goal is a demonstrable federated learning system where, for example, communication costs are drastically lower than central training and privacy metrics meet regulatory guidelines. By this stage, we expect to report quantitative metrics such as the bytes communicated per training round, the convergence time to reach a target performance, and the achieved privacy level alongside any trade-offs in accuracy.

Beyond 2 years, the vision is to evolve the intelligent mobility system from advanced prototyping to real-world deployment, while pushing the boundaries of autonomy, collaboration, and security:

**Real-World Deployment & Continuous Learning:** Transition the LLM-integrated autonomous driving system from the digital twin and test fleets to real-world operating environments. This will likely involve phased deployments in controlled settings before broader rollout. A key long-term goal is to enable continuous learning and adaptation: as vehicles encounter new situations on the road, they will leverage federated learning to continuously improve the shared models without requiring downtime or centralized data collection. The digital twin orchestration platform will play a pivotal role in this deployment. By mirroring each vehicle with a live digital twin, the system can perform real-time shadow evaluations of novel scenarios in the virtual space before the vehicle takes action, thus bridging the gap between theoretical modeling and safe real-world behavior. The long-term outcome is a fleet of LLM-enabled autonomous agents that improve over time, providing evidence of learning in the wild while new knowledge is instantly shared across the fleet through federated updates.

**Advanced LLM–Agent Capabilities:** In the long run, the integration of LLMs will be extended to more complex and critical driving tasks. We envision LLMs handling higher-level reasoning such as cooperative maneuvers between vehicles, strategic route planning under congested or hazardous conditions, and semantic understanding of traffic rules and human gestures. The agent architecture will evolve into a hierarchical model: the LLM acting as a high-level cognitive layer for reasoning and communication, and the low-level controllers handling execution. This arrangement will be refined to ensure real-time performance and fail-safe operation. Moreover, the LLMs could be specialized or fine-tuned on mobility-specific corpora by this stage, giving them a richer understanding of traffic narratives and regulations. By more than 2 years, we aim for the LLM-augmented system to be capable of complex multi-agent interactions, which would be validated first in the simulator and then in field trials. Throughout this, rigorous testing and verification will be conducted to certify that the LLM's behavior meets safety and reliability standards required for autonomous driving deployments.

**Federated Learning 2.0 – City-Scale and Fully Privacy-Preserving:** As the number of deployed autonomous vehicles grows, the federated learning framework will be scaled to city-scale fleets

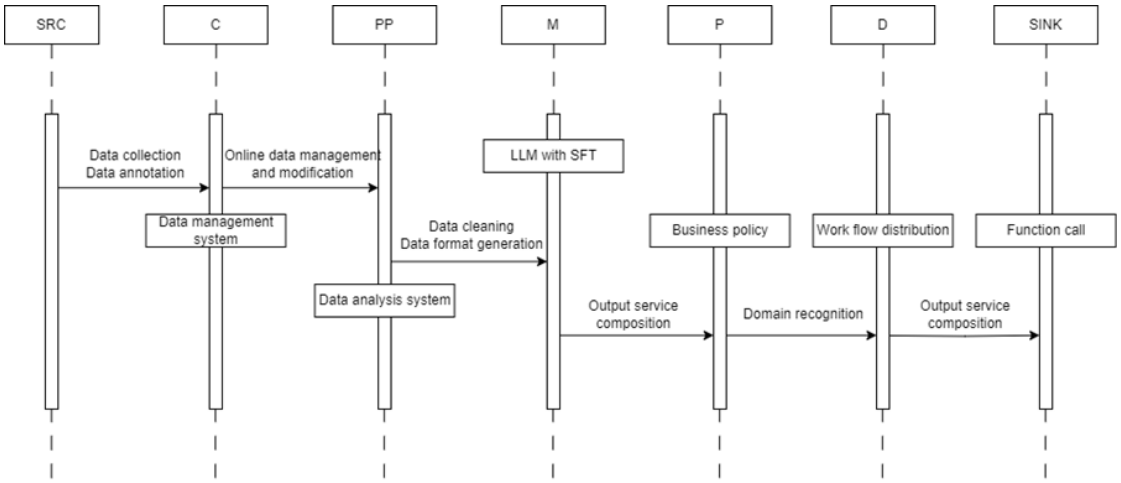
and beyond. The long-term goal is a globally distributed learning network where thousands of vehicles and infrastructure units collaborate to train robust AI models for perception and decision-making. Emphasis will be on maintaining system scalability, fairness, and rigorous security despite the scale. Techniques such as hierarchical federated learning may be employed to manage communication loads. We will target stringent privacy guarantees, aiming for strong differential privacy levels and incorporating multi-faceted privacy protection: for instance, combining differential privacy with secure aggregation and encryption so that even if some infrastructure is compromised, individual vehicle data remains protected. We will also define clear privacy metrics to monitor and ensure compliance with data protection regulations. In parallel, research into communication-efficient FL will continue. Ultimately, by leveraging these advanced federated learning approaches, the system aims to learn from vast distributed data in a privacy-preserving manner, resulting in highly accurate models that no single entity could have trained alone. Success will be measured by demonstrating that a model trained in this federated network achieves performance on par with a hypothetically centrally-trained model, as well as by meeting targets for communication efficiency and privacy.

**V2X Security and System Robustness:** A major long-term focus will be hardening the entire system against security threats and failure modes, especially as it transitions to real-world critical infrastructure. This entails enhancing V2X system security on multiple fronts. All vehicle-to-infrastructure and vehicle-to-vehicle communications will employ state-of-the-art encryption, authentication, and integrity-checking in line with automotive standards. For instance, messages exchanged will be signed and verified to prevent spoofing, and sensitive updates in the federated learning process will utilize secure aggregation to thwart eavesdroppers. The digital twin platform can be utilized to simulate cyber-attacks to test the system's resilience and to refine intrusion detection algorithms. Additionally, the interplay of LLMs introduces new considerations for robustness: measures will be taken to prevent and mitigate issues like malicious or ambiguous inputs to the LLM, and to ensure the LLM's outputs do not compromise safety. By incorporating these defensive techniques and continuously validating them, the system is expected to maintain a high level of trust and safety. The end-state vision is an intelligent mobility network where autonomous agents cooperate seamlessly using learned intelligence, and where both data privacy and communication security are upheld by design. This will instill confidence that the transition from theoretical models to real-world operation does not introduce vulnerabilities, thus paving the way for widespread adoption of the technology.

## 3 Use Case Requirements

- **REQ-01**: *It is mandatory* that the system operates on a standard 4G network, as the transmitted data consists of lightweight text tokens; however, *it is also critical* that network coverage exceeds 95% to avoid service disruptions and ensure a continuous user experience.
- **REQ-02**: *It is critical* that the vehicle is equipped with a Service-Oriented Architecture (SOA), with in-vehicle functions modularized into atomic services to enable seamless orchestration.
- **REQ-03**: *It is mandatory* that the system supports API interfaces for large language model cloud services from certified providers, ensuring legal compliance, high stability, and scalability to handle varying concurrency levels.

## 4 Sequence Diagram




## 5 References

[1] Douyin. "Video: Changan Automobile – AI for Good." Accessed June 19, 2025. https://www.douyin.com/video/7510067871398448443.

[2] Chongqing Changan Automobile Co., Ltd. AI for Good. GitHub. Accessed June 19, 2025. https://github.com/Chongqing-Changan-Automobile-Co-Ltd/AI_for_good.

[3] Chongqing Changan Automobile Co., Ltd. Changan Qiyuan. GitHub. Accessed June 19, 2025. https://github.com/Chongqing-Changan-Automobile-Co-Ltd/AI_for_good/blob/main/Changan_Qiyuan.pdf.

[4] Chongqing Changan Automobile Co., Ltd. Seizing the Smart-Tech High Ground: The Confidence Behind Changan's Rise on the Global Stage. GitHub. Accessed June 19, 2025. https://github.com/Chongqing-Changan-Automobile-Co-Ltd/AI_for_good/blob/main/Seizing_the_Smart-Tech_High_Ground_The_Confidence_Behind_Changan%E2%80%99s_Rise_on_the_Global_Stage.pdf.

# Use Case 6: Enhancing Road Coordination Through Vehicle-to-Vehicle (V2V) Communication

**Organization:** OptM Media Solutions Private Ltd

**Country**: India

**Contact Person(s)**:

Umesh Tallam, CE0, umesh.tallam@optmsol.com, 9886722627

Chetan Mohan, Data Engineering, chetan.mohan@optmsol.com, 78299442058

## 1 Use Case Summary Table

| Item | Details |
|---|---|
| Category | Vehicle to Vehicle(V2V) |
| Problem Addressed | Narrow roads cause vehicle congestion and accidents due to poor coordination and limited maneuvering space. |
| Key Aspects of Solution | AI-driven V2V communication enables real-time vehicle coordination using sensor data and reinforcement learning. |
| Technology Keywords | V2V communication, AI, IoT, edge computing, autonomous systems, real-time decision-making, GPS. |
| Data Availability | |
| Metadata (Type of Data) | Sensor data, GPS coordinates, road width, weather conditions, vehicle size and movement information. |
| Model Training and Fine- | Reinforcement learning trains AI models for real-time decisions based on live road condition data. |
| Testbeds or Pilot Deployments | Pilot tests on India's rural and semi-urban roads using controlled environments and mixed vehicle types. |

# 2 Use Case Description

## 2.1 Description

**Context & Background**:

India's road infrastructure, particularly in rural and semi-urban areas, is often characterized by narrow stretches that are not adequately designed to handle the increasing volume and diversity of vehicular traffic. These roads frequently lack proper lane markings, shoulders, and designated pedestrian pathways, making it difficult for vehicles to maneuver safely. In such conditions, two-way traffic becomes a challenge, especially when larger vehicles like buses, trucks, or tractors encounter smaller vehicles or pedestrians. This lack of space often results in traffic bottlenecks, slow movement, and increased chances of head-on collisions or sideswiping accidents. Additionally, the absence of proper traffic management and enforcement further exacerbates the situation, leading to frequent congestion, delays, and safety hazards for all road users.

Objectives & Aims:

The primary objective of this initiative is to enhance coordination among vehicles through the integration of Vehicle-to-Vehicle (V2V) communication and Artificial Intelligence (AI) technologies. By enabling real-time data exchange between vehicles, the system aims to improve situational awareness, especially on narrow and congested roads. AI will play a crucial role in analysing traffic patterns, predicting potential hazards, and providing automated alerts or suggestions to drivers, thereby reducing the risk of accidents caused by human error. Additionally, coordinated vehicle movement facilitated by V2V communication is expected to ease traffic congestion and promote smoother traffic flow. The ultimate goal is to create a smart and adaptive transportation network that ensures greater road safety, efficiency, and a more reliable driving experience for all road users.

**Technological Approach:**

The proposed solution relies on a combination of Internet of Things (IoT)-based Vehicle-to-Vehicle (V2V) communication systems and advanced Artificial Intelligence (AI) models to enhance road safety and traffic management. Through real-time communication enabled by IoT sensors and devices embedded in vehicles, essential data such as speed, location, road conditions, and vehicle intentions are continuously shared among nearby vehicles. This interconnected network allows for instantaneous updates and alerts, ensuring vehicles can respond proactively to dynamic road scenarios. Complementing this, AI models– particularly those trained using reinforcement learning techniques–are employed to analyze complex traffic environments and make intelligent decisions. These models learn optimal driving behaviors through continuous feedback, enabling vehicles to perform safe and efficient maneuvers such as lane merging, overtaking, or yielding in real time. Together, these technologies work in tandem to build a responsive, adaptive, and intelligent driving ecosystem.

**Expected Impact & Benefits:**

The implementation of V2V communication and AI-driven decision-making is expected to significantly improve navigation and safety on narrow and congested roads. By enabling

vehicles to coordinate their movements intelligently, the system will help prevent accidents that typically occur due to poor visibility, misjudgement, or lack of space. Traffic congestion is also expected to reduce, as vehicles can respond dynamically to real-time road conditions and optimize their paths accordingly. This enhanced traffic flow not only contributes to shorter travel times but also leads to reduced fuel consumption and lower emissions, supporting environmental sustainability. Overall, the integration of these smart technologies aims to create safer, more efficient, and eco-friendly transportation systems, particularly in regions with challenging road infrastructure.

The proposed system must meet several technical, operational, and functional requirements to ensure its effectiveness and reliability in real-world scenarios. First, it must support real-time communication within a minimum range of 100 meters (REQ-01), allowing vehicles to exchange critical data promptly. Each vehicle should be equipped with GPS and onboard sensors to enable accurate positioning and environmental awareness (REQ-02). The communication protocols used must adhere to established global V2V standards to ensure interoperability and future scalability (REQ-03). To facilitate timely decision-making, AI models integrated into the system must process data and generate responses within 500 milliseconds (REQ-04). Given the variability of Indian road and weather conditions, the system must maintain reliable performance in fog, rain, heat, and other environmental factors (REQ-05). Lastly, robust data privacy and cybersecurity measures are essential to protect sensitive vehicle and user information from unauthorized access or misuse (REQ-06). Meeting these requirements is vital for the safe, efficient, and trustworthy deployment of the V2V-AI solution.

### Purpose of the Flow

This likely represents a **collaborative driving or autonomous vehicle decisionmaking** process in scenarios where two vehicles need to negotiate space (like a single-lane road or obstacle avoidance). It ensures **safe and smooth navigation** without collisions or confusion by utilizing **system-mediated communication and decision-making**.

## DATA DESCRITION

1. Domain Context

- The dataset is from an IoT-based network environment.
- Focused on cybersecurity, specifically network intrusion detection.
- Each record corresponds to a **network flow/session** (typically unidirectional traffic between endpoints).

2. Dataset Scale

- Contains **~46 million** total records.
- Organized into **163 CSV batch files** for efficient processing and training.

Feature Overview (Total: 46 Features + 1 Label)

The dataset comprises detailed network flow records from an IoT environment, structured with traffic durations, protocol flags, application indicators, and statistical descriptors. It includes binary protocol usage, transmission rates, and geometric metrics for modelling. Designed for intrusion detection and cyberattack classification, it supports both classical and deep learning approaches.

**Partners:** NXP : We are using LiDAR and Radar Systems and SiMa.ai :We are using ECU board

**Impact on Intelligent Transport:**

The proposed intelligent traffic systems play a vital role in advancing sustainable and innovative transport solutions. By enhancing road infrastructure and integrating advanced technologies, these systems improve overall safety and reliability across urban and interurban transport networks.

Intelligent vehicle coordination enables smoother traffic flow and reduces congestion, contributing to more efficient and sustainable urban mobility. As vehicles and infrastructure communicate in real time, travel becomes more predictable, reducing delays and improving commuter experience.

Additionally, optimizing traffic flow through smart systems directly lowers fuel consumption and vehicle emissions. This not only supports environmental protection efforts but also reduces the carbon footprint of transportation, aligning with broader goals for climate resilience.

By combining innovation, infrastructure development, and environmentally responsible practices, intelligent transport systems lay the foundation for cleaner, safer, and more livable cities.

## 2.2 Benefits of use case

The proposed intelligent traffic systems play a vital role in advancing sustainable and innovative transport solutions. By enhancing road infrastructure and integrating advanced technologies, these systems improve overall safety and reliability across urban and interurban transport networks.

Intelligent vehicle coordination enables smoother traffic flow and reduces congestion, contributing to more efficient and sustainable urban mobility. As vehicles and infrastructure communicate in real time, travel becomes more predictable, reducing delays and improving commuter experience.

Additionally, optimizing traffic flow through smart systems directly lowers fuel consumption and vehicle emissions. This not only supports environmental protection efforts but also reduces the carbon footprint of transportation, aligning with broader goals for climate resilience.

By combining innovation, infrastructure development, and environmentally responsible practices, intelligent transport systems lay the foundation for cleaner, safer, and more livable cities.

## 2.3 Future work

Planned Enhancements and Developments:

- Expand vehicle coordination capabilities to operate effectively in dense urban areas and complex traffic environments.
- Integrate **Vehicle-to-Infrastructure (V2I)** communication with roadside units, traffic signals, and environmental sensors.

- Conduct full-scale pilot deployments in real-world, multi-vehicle traffic scenarios to evaluate system performance and scalability.

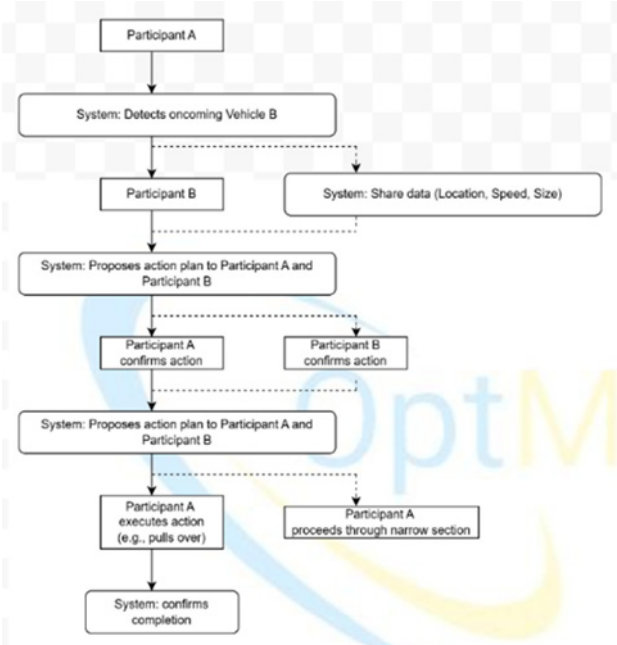
Technology Upgrades and Data Needs:

- Upgrade **edge computing nodes** to enable faster, more localized processing and real-time decision-making.
- Enhance AI model robustness by training with **context-aware road data** that covers diverse weather conditions and terrain types.
- Ensure seamless interaction between V2V and V2I systems through synchronized data exchange protocols.
- Collect and process high-resolution traffic, sensor, and environmental data for improved AI learning and adaptability.

## 3 Use Case Requirements

- **REQ-01:** It is critical that the system supports **real-time communication within a minimum 100-meter range** to enable vehicles to exchange critical data quickly and reliably.
- **REQ-02**: It is mandatory that **vehicles are equipped with GPS and onboard sensors** to ensure accurate tracking, positioning, and real-time environmental awareness.
- **REQ-03**: It is critical that **communication protocols align with global Vehicle-to-Vehicle (V2V) standards**, ensuring system interoperability, compatibility, and scalability across different regions and manufacturers.
- **REQ-04**: It is mandatory that **AI decision-making processes are completed within 500 milliseconds** to support rapid, time-sensitive actions and ensure safety in dynamic driving environments.
- **REQ-05**: It is critical that the system functions **reliably in diverse weather conditions**, including fog, rain, and heat, to maintain consistent performance on Indian roads and other challenging environments.
- **REQ-06: It is mandatory** to implement **strong data privacy and cybersecurity measures** to safeguard sensitive vehicle and user data from unauthorized access, breaches, or misuse.

## 4 Sequence Diagram




This sequence diagram illustrates a cooperative decision-making process between two participants (A and B) with the help of systems likely in an autonomous or semi-autonomous vehicle scenario. Here is the breakdown of the system

1. Initial Detection

- Participant A is in a scenario where the system detects an oncoming vehicle (Vehicle B).

2. Data Sharing

- The system shares relevant information (like location, speed, and size) with Participant B.
- This sharing likely occurs over a network to help both participants understand the traffic situation.

## 5 References

[1] Fountain, Ernest, Lo'ai Tawalbeh, and Jeong Yang. "Predicting Volume of Vehicular Traffic Using Machine Learning." Issues in Information Systems 21, no. 3 (2020). https://openurl.ebsco.com/EPDB%3Agcd%3A11%3A2148907/detailv2?sid=ebsco%3Aplink%3Ascholar&id=ebsco%3Agcd%3A145743834&crl=c&link_origin=scholar.google.com.

[2] Ahrems, J. "Development of V2V and GPS Based Collision Warning Algorithm for Uncontrolled Intersections." In 2017 Electronics, Palanga, Lithuania, 2017, 1–4. https://doi.org/10.1109/ELECTRONICS.2017.7995227.

[3] Nguyen, T., N. T. Le, and Y. M. Jang. "Asynchronous Scheme for Optical Camera Communication-Based Infrastructure-to-Vehicle." International Journal of Distributed Sensor Networks (2015). https://journals.sagepub.com/doi/full/10.1155/2015/908139.

[4] 3GPP. NR; User Equipment (UE) Radio Transmission and Reception – Part 4: Performance Requirements. TS 38.101-4, Version 17.4.0 Release 17. 2022.

[5] 3GPP. Study on Evaluation Methodology of New Vehicle-to-Everything (V2X) Use Cases for LTE and NR. TR 37.885, Version 15.3.0. June 2019. https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3209.

# Use Case 7: AI-Driven 6G-Enabled V2X Communications for Smart and Sustainable Mobility

**Organization:** AI4Africa Research Group

**Country:** Nigeria

**Contact Person(s)**:

Emmanuel Aaron (aaronemmanuel054@gmail.com, +2348077200689),

Dr. Houda Chihi (houda.chihi@supcom.tn),

Blessed Guda (gudablessed@gmail.com),

Emmanuella Sule (suleemmanuella@yahoo.com),

Chidi Ebube (chidizack24@gmail.com),

Emmanuel Ani (ani.mlengineer@outlook.com),

Okafor Miracle Uche (okaformiracle212@gmail.com)

## 1 Use Case Summary Table

| Item | Details |
|---|---|
| Category | Intelligent transport, v2x, AI |
| Problem Addressed | Network interference and mobility, resource allocation among vehicular nodes |
| Key Aspects of Solution | Integer Linear Programming (ILP) is employed for network resource allocation in environments with fixed conditions, while Deep Reinforcement Learning (DRL) is utilized for adaptive resource allocation in vehicular networks with varying mobility patterns [1]. |
| Technology Keywords | Autonomous Vehicles, V2X, network resource allocation, vehicular mobility. |
| Data Availability | Data Generation is extracted using public access for Open Street Map[2] for DRL training. |
| Metadata (Type of Data) | Matrices (for the network states), numeric data from RSU logs for latency, load, or network fluctuation, LiDAR or sensor readings from vehicles. |

(continued)

| Item | Details |
|---|---|
| Model Training and Fine-Tuning | Combining Integer Linear Programming and Deep Reinforcement Learning (Proximal Policy Optimization). |
| Testbeds or Pilot Deployments | No Deployments at the moment, but the intended toolkits include the Sumo-Gym toolkit[3]. |
| Code repositories | Code Repo and model (in progress)[4]. |

# 2 Use Case Description

## 2.1 Description

The rapid growth of autonomous and connected vehicles presents enormous challenges to road safety, network scalability, and environmental sustainability. Existing V2X communication systems are impacted by network interference, mobility, and latency, together with connectivity challenges, especially in high traffic density in urban areas, leading to handover enhancement.

The proposed use case explores an AI-driven 6G V2X architecture[5] that is designed to enable scalable, reliable, and energy-efficient vehicular communication. Through the exploitation of Deep Reinforcement Learning (DRL)–i.e., Proximal Policy Optimization (PPO) algorithm–the system dynamically allocates vehicles to Roadside Units (RSUs), offering better optimization in the concern of energy, latency, and quality of service. A custom simulation environment built on Gymnasium trains the DRL agent to respond in real-time against uncertain traffic and network states. Integer Linear Programming (ILP), as conventionally used, is taken as a benchmark to check the performance of DRL. Let RSU correspond to the agent who takes the decision following an interaction with the vehicular environment.

When selecting an appropriate RSU for a moving vehicle, the system considers several important factors to provide the best performance. These include signal strength (RSSI or SINR), distance between the vehicle and RSU, available resource blocks, RSU load at the moment, processing capacity, and latency guarantees. The DRL agent (deployed in the RSU) learns from these factors, which the RSU provides the best trade-off between energy efficiency and quality of service. This multi-objective decision-making is tasked with the responsibility of providing a steady, low-latency link for each vehicle without compromising the entire network load's balance. The idea is to deploy the trained agent on the RSU or Multi-access Edge Computing infrastructure. The Federated learning is centralized, with every new scenario learned updated on a centralized system.

The entire situation—RSU rollout, vehicle movement, and criteria for decisions—can be fully simulated. Cars can be simulated with realistic or random movement in urban or highway settings. RSUs are randomly distributed on the map. The environment calculates useful metrics like SINR, proximity, and availability of RSUs at each timestep so that the DRL agent can decide optimal RSU allocation. Gymnasium, SUMO, or OMNeT++ are some of the tools used to simulate true behaviour so that AI models can be trained and tested robustly in a safe, controlled virtual environment. The integration between the Gymnasium environment and SUMO is possible through the use of already implemented Python toolkits, *Sumo-gym* or *Sumo-rl*. Using this toolkit, the agent can be trained and previewed in a 2D simulation concurrently.

The optimal RSU choice depends on some factors like the coverage radius (distance), the latency, the speed of the vehicle, e.t.c. The reward function for the model is shown in equation (1), where:

- $R = (1 - 0.1 \times (1 + \ell_j/N))$: Reward for connecting to an RSU with light load and strong signal (**latency**)
- -1: Penalty if the vehicle is out of the RSU range or has a weak signal
- $-\lambda \times (d_{ij} / v)$: **Energy cost** penalty based on distance

Where,

- V: Set of vehicles indexed by i = 1, 2, ..., N
- R: the corresponding reward.
- $d_{ij}$: Distance from vehicle i to RSU j
- $s_{ij}$: Signal strength between vehicle i and RSU j
- $\ell j$: Current load on RSU j
- $v$: RSU coverage radius

If the agent selects an RSU that does not meet the criteria for optimal distance and load, a flat penalty of −1 is applied. Additional penalties are also incorporated to account for energy consumption (based on distance) and the probability of an imminent handover, which is influenced by the vehicle's velocity. These factors are captured through their respective cost functions.

This use case is currently in the model training and simulation phase, and so the results are yet to be produced. While there are other scenarios where this methodology can be applied—e.g., semantic communication and predictive analytics to improve vehicle decision-making or forecast road hazards, and AI-based traffic management through strategies like platooning and adaptive signal timing—these fall outside the scope of this use case. However, they remain feasible through further research and the training of time series models, such as LSTMs, on open datasets to help develop more efficient traffic systems.

In total, AI-enabled 6G V2X systems offer a groundbreaking solution to the current technological limitations, rendering transportation networks safe, efficient, and sustainable for future autonomous mobility.

### Category – Intelligent Transport

Intelligent transport systems powered by AI-enabled 6G-V2X networks directly strengthen mobility infrastructure, delivering secure, friction-free communication for autonomous and connected vehicles. Predictive maintenance, real-time tracking, and dynamic route optimisation make roads and highways safer, more efficient, and more productive while encouraging continuous innovation across the transport sector.

These capabilities also underpin the development of sustainable, liveable cities. AI-driven traffic management eases congestion, improves public-transport reliability, and keeps emergency vehicles moving, all of which enhance urban air quality and quality of life for residents.

Finally, streamlined traffic flows and coordinated driving modes such as platooning cut fuel consumption and tail-pipe emissions, supporting cleaner, greener environments without compromising the reliability or security of road transport in either urban or highway settings.

## 2.2 Benefits of use case

Intelligent transport systems powered by AI-enabled 6G-V2X networks directly strengthen mobility infrastructure, delivering secure, friction-free communication for autonomous and connected vehicles. Predictive maintenance, real-time tracking, and dynamic route optimisation make roads and highways safer, more efficient, and more productive while encouraging continuous innovation across the transport sector.

These capabilities also underpin the development of sustainable, liveable cities. AI-driven traffic management eases congestion, improves public-transport reliability, and keeps emergency vehicles moving, all of which enhance urban air quality and quality of life for residents.

Finally, streamlined traffic flows and coordinated driving modes such as platooning cut fuel consumption and tail-pipe emissions, supporting cleaner, greener environments without compromising the reliability or security of road transport in either urban or highway settings.

## 2.3 Future work

Future research in this use case will focus on advancing key areas that will shape the evolution of AI-based 6G V2X communications. The foremost priority will be the expansion of data collection from real-world traffic scenarios, aiming to build a comprehensive, diverse dataset that reflects various vehicular interactions, road types, and urban mobility patterns. This will involve strategic collaborations with smart city initiatives, transportation departments, and automotive industry partners to ensure access to high-quality, live data.

Following the data acquisition phase, a proof-of-concept system will be developed and tested within highly regulated V2X testbed environments. These testbeds will allow for rigorous evaluation of AI models under real-world conditions, assessing their ability to deliver accurate predictive analysis, vehicle coordination, and adaptive communication strategies in dynamic road settings.

Model development will play a central role, with a focus on designing advanced AI algorithms for decision-making, anomaly detection, and traffic flow optimization. This will include the implementation of deep learning models for real-time processing and federated learning techniques to enable decentralized, privacy-preserving computation across networked vehicles and edge nodes.

Additionally, collaboration with the International Telecommunication Union (ITU) and other regulatory bodies will be pursued to contribute toward the standardization of AI-driven 6G V2X protocols, promoting global interoperability and adoption.

A digital twin-based simulation framework will also be established to conduct large-scale testing of AI-powered V2X systems in virtual environments before real-world deployment.

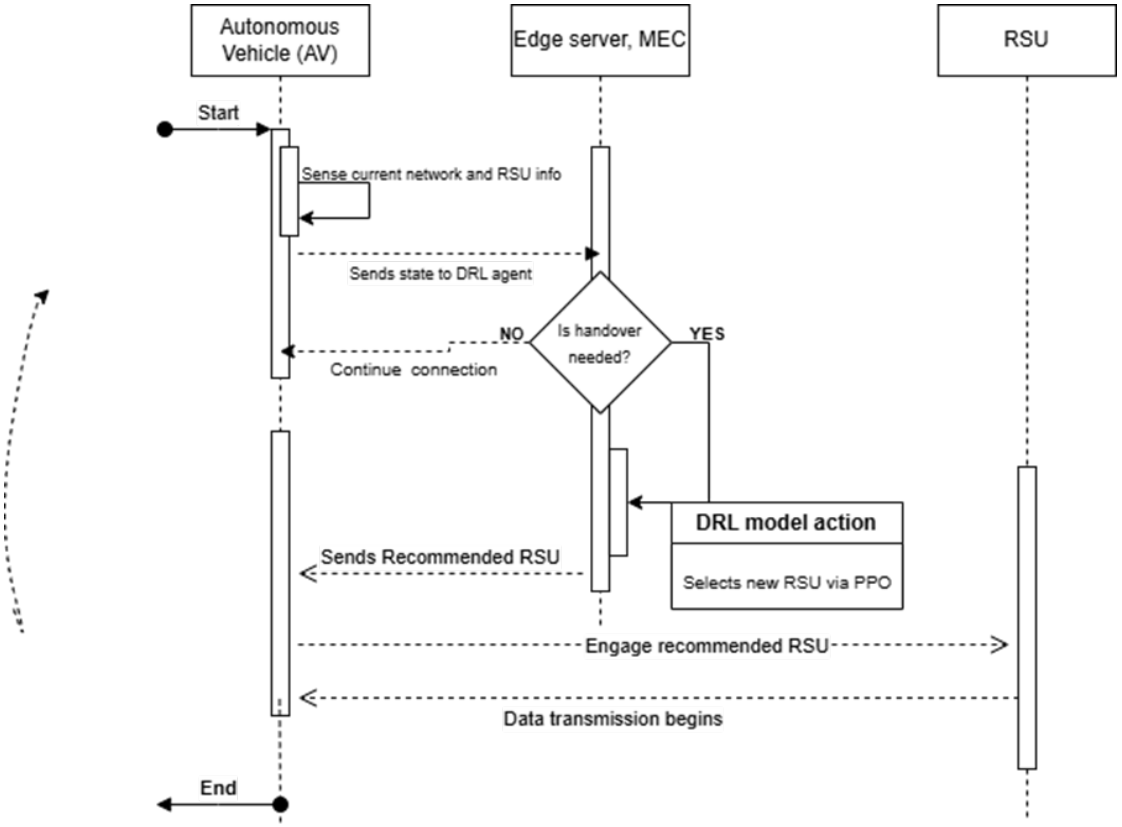
Ultimately, this work aims to build the foundation for fully autonomous, intelligent V2X communication systems that support safer, more efficient, and sustainable transportation

infrastructures. The long-term vision is to seamlessly integrate AI-enabled 6G V2X technologies into smart cities, enhancing mobility while minimizing environmental impact.

## 3 Use Case Requirements

- **REQ-01**: *It is mandatory* that the system supports wireless infrastructure with ultra-low latency and high data throughput to enable real-time V2X communication[5].
- **REQ-02**: *It is critical* to implement a Deep Reinforcement Learning (DRL) algorithm—specifically Proximal Policy Optimization (PPO)—to dynamically manage vehicle-to-RSU assignments.
- **REQ-03**: *It is critical* that the DRL agent adapts in real time to network changes such as traffic density, RSU load, and vehicle mobility.
- **REQ-04**: *It is mandatory* to incorporate a retraining mechanism that updates the DRL model based on evolving traffic patterns and urban layouts.
- **REQ-05**: *It is mandatory* that each vehicle be assigned to exactly one RSU at any given time, ensuring a valid and exclusive communication link.
- **REQ-06**: *It is critical* that the system minimizes energy consumption across RSUs while still satisfying quality of service (QoS) constraints such as latency and bandwidth.
- **REQ-07**: *It is critical* for the DRL agent to continuously learn from environmental feedback to enhance future decision-making and improve resource efficiency.

## 4 Sequence Diagram



## 5 References


[1] Li, P., Z. Xiao, H. Gao, X. Wang, and Y. Wang. "Reinforcement Learning Based Edge-End Collaboration for Multi-Task Scheduling in 6G Enabled Intelligent Autonomous Transport Systems." IEEE Transactions on Intelligent Transportation Systems (2024). https://doi.org/10.1109/TITS.2024.3525356.

[2] OpenStreetMap contributors. OpenStreetMap. https://www.openstreetmap.org/#map=4/21.84/82.79.

[3] "sumo-gym." PyPI. Accessed June 19, 2025. https://pypi.org/project/sumo-gym/.

[4] "6G V2X Research Folder." Google Drive. Accessed June 19, 2025. https://drive.google.com/drive/folders/1YfdElxsONthgZrjw7F3-3E0hWyBVsEA.

[5] Aung, P. S., L. X. Nguyen, Y. K. Tun, Z. Han, and C. S. Hong. "Deep Reinforcement Learning-Based Joint Spectrum Allocation and Configuration Design for STAR-RIS-Assisted V2X Communications." IEEE Internet of Things Journal 11, no. 7 (April 1, 2024): 11298–11311. https://doi.org/10.1109/JIOT.2023.3329893.

[6] Paulchen2713. V2X-RA-DRL. GitHub. https://github.com/paulchen2713/V2X-RA-DRL.

[7] Zhang, X., M. Peng, S. Yan, and Y. Sun. "Deep-Reinforcement-Learning-Based Mode Selection and Resource Allocation for Cellular V2X Communications." IEEE Internet of Things Journal 7, no. 7 (July 2020): 6380–6391. https://doi.org/10.1109/JIOT.2019.2962715.

[8] Annu, and P. Rajalakshmi. "Towards 6G V2X Sidelink: Survey of Resource Allocation—Mathematical Formulations, Challenges, and Proposed Solutions." IEEE Open Journal of Vehicular Technology 5 (2024): 344–383. https://doi.org/10.1109/OJVT.2024.3368240.

# Use Case 8: AI for Intrusion in Vehicle-to-Vehicle (V2V) Communication

**Organization:** AI4Africa Research Group

**Country:** Nigeria

**Contact Person(s)**:

Chidi Ebube (chidizack24@gmail.com),

Dr Houda Chihi (houda.chihi@supcom.tn),

Blessed Guda (gudablessed@gmail.com),

Emmanual Aaron (aaronemmanuel054@gmail.com),

Emmanuella Sule (suleemmanuella@yahoo.com),

Emmanuel Ani (ani.mlengineer@outlook.com),

Okafor Miracle Uche (okaformiracle212@gmail.com)

## 1 Use Case Summary Table

| Domain | Intelligent Transport |
|---|---|
| Problem to be Addressed | Detect and prevent intrusions or attacks (e.g., spoofing, jamming, unauthorized data injection) within V2V communication systems to ensure vehicle safety and coordination. |
| Key Aspect of the Solution | Deep learning-based intrusion detection system (IDS) using LSTM for sequential anomaly detection and Autoencoders for outlier detection in V2V communication. |
| Technology Keywords | Deep Learning, LSTM, Autoencoders, Intrusion Detection System (IDS), Vehicle-to-Vehicle (V2V), OBUs (Onboard Units), Anomaly Detection, Edge Computing. |
| Data Availability | V2X-Sim Dataset [5]<br>Network Intrusion Dataset [6] |
| Metadata (Type of Data) | Vehicle communication logs, sensor data, traffic information (speed, location, braking, warnings), labeled data (normal/abnormal communication patterns), timestamps, event types (e.g., collision warnings). |

(continued)

| Domain | Intelligent Transport |
|---|---|
| Model Training and Fine-Tuning | - **LSTM**: Detects anomalies in sequential V2V communication data.<br>- **Autoencoders**: Learns normal traffic patterns and flag outliers as potential intrusions.<br>- **Supervised & Unsupervised Learning**: Uses labeled attack data where available and unsupervised methods where labels are missing. |
| Testbeds or Pilot Deployments | Pilot deployment on **OBUs in real-world vehicles** or **simulation environments** to test the IDS's real-time performance and accuracy in detecting intrusions. Can use vehicle communication simulators (V2X-Sim) or controlled traffic scenarios. |

# 2 Use Case Description

## 2.1 Description

The **Internet of Things (IoT)** has transformed various industries over the years, and now a new emerging field in transportation safety, known as the **Internet of Vehicles (IoV)**, is gaining prominence. While the core technology behind IoV is **Vehicle-to-Everything (V2X)** communication, where vehicles interact with their surroundings, the necessity for this technology arises from a pressing global concern—**road safety**.

Every year, a significant number of fatal road accidents occur, making transportation safety a top priority. Vehicle-to-Vehicle (V2V) communication is a crucial subset of V2X that enables vehicles to exchange real-time information to prevent collisions, coordinate movement, and enhance road safety. However, as with any technological advancement, new vulnerabilities and security threats emerge. The dataset employed in this research is the V2X-Sim 2.0 dataset, a multi-sensor, annotated dataset specifically developed for research in autonomous driving and vehicle-to-everything (V2X) communication. It follows the same schema as the nuScenes dataset, ensuring a well-structured and comprehensive format for training and evaluating perception models in V2V environments.

One of the most significant challenges in V2V communication is ensuring the security of transmitted data against cyber threats. Malicious activities such as false data injection, spoofing, Sybil attacks, and denial-of-service (DoS) attacks can compromise network integrity. If false information is introduced into the system, it can mislead vehicles and lead to hazardous situations, potentially endangering lives. This highlights the importance of robust Intrusion Detection Systems (IDS) tailored to the unique demands of V2V networks.

For evaluating our IDS, we will use standard classification metrics, including Accuracy, Precision, Recall, and F1-Score.

In addition to V2X-Sim, we will utilize the CIC-IDS-2017 dataset, a widely used benchmark for network intrusion detection. This dual-dataset approach supports both perception-level data from vehicle sensors and network-level traffic data for intrusion analysis.

We employ Long Short-Term Memory (LSTM) networks and Autoencoders as our primary deep learning models:

1. LSTM is well-suited for modeling time-series data, capturing sequential patterns in vehicular communication.
2. Autoencoders, as an unsupervised learning method, are effective for anomaly detection, learning the normal behavior of network transmissions, and identifying outliers or deviations.

This deep learning-based approach offers a clear advantage over traditional signature-based or rule-based IDS methods. Traditional IDSs are typically designed for relatively static networks and rely on known attack patterns, making them ineffective against zero-day attacks. In contrast, our approach is dynamic, adaptive, and better suited for the mobile, real-time nature of vehicular communication.

We adopt a dual-dataset strategy in our data pipeline. Initially, the V2X-Sim dataset will be used in a simulation environment to extract multi-agent sensor recordings from multiple vehicles. Simultaneously, we will leverage the CIC-IDS-2017 dataset for evaluating network-level intrusion detection.

Intrusion Detection Systems (IDSs) and Intrusion Prevention Systems (IPSs) remain critical defense mechanisms against the growing sophistication of network attacks. However, anomaly-based IDSs often face challenges due to the limited availability of reliable, diverse datasets, which affects consistent and accurate performance evaluation.

To overcome this limitation, our primary testbed will be the open-source V2X-Sim platform, enabling low-latency simulation and real-time data extraction. This setup allows our system to remain flexible and adaptive in various scenarios. Additionally, to further enrich the training data, we plan to use Generative Adversarial Networks (GANs) to create synthetic attack scenarios, enhancing our model's robustness against unseen threats.

This research aims to tackle these security challenges by developing an intrusion detection system (IDS) powered by deep learning, specifically LSTM (Long Short-Term Memory) and Autoencoders. The goal is to detect anomalous or malicious V2V messages in real-time, ensuring that only authentic and trustworthy information is shared among vehicles, thereby enhancing the security and reliability of autonomous and connected transportation systems.

## 2.2 Benefits of use case

This research strengthens the foundation of intelligent transport systems by focusing on cybersecurity in vehicle-to-vehicle (V2V) communication, an essential component of modern connected infrastructure. Intelligent transport networks rely heavily on accurate and timely data exchange between vehicles, but they face serious threats from cyberattacks such as false data injection, GPS spoofing, and denial-of-service (DoS) attacks. These can lead to major disruptions, including traffic accidents and systemic failures. By addressing these vulnerabilities, the research helps safeguard critical infrastructure and ensure the safe operation of smart transportation systems.

A secure V2V communication network is also vital for sustainable urban mobility. Manipulated or hacked data can trigger false alerts, misdirect traffic, or cause unnecessary congestion—all of which compromise public safety and efficiency. This work ensures that connected vehicles

share only trustworthy, verified information, thus improving traffic flow, reducing the likelihood of accidents, and enabling safer autonomous driving environments.

Moreover, the research supports broader efforts to strengthen digital governance in transportation. Cyber threats in this space not only affect physical safety but also challenge public trust and institutional control. By developing real-time cyberattack detection and mitigation methods, this work contributes to the protection of digital infrastructure and aligns with global efforts to combat cybercrime in critical systems, promoting accountability, resilience, and long-term confidence in smart mobility solutions.

### 2.3 Future work

As V2X communication and AI-based intrusion detection systems continue to evolve, there are several areas where further improvement if provided with the scholarship

Firstly, since this research focuses on V2V security, it is crucial to validate the model in real-world scenarios. Deploying the Intrusion Detection System (IDS) on On-Board Units (OBUs) will require specialized hardware and field testing to ensure effectiveness in dynamic vehicular environments.

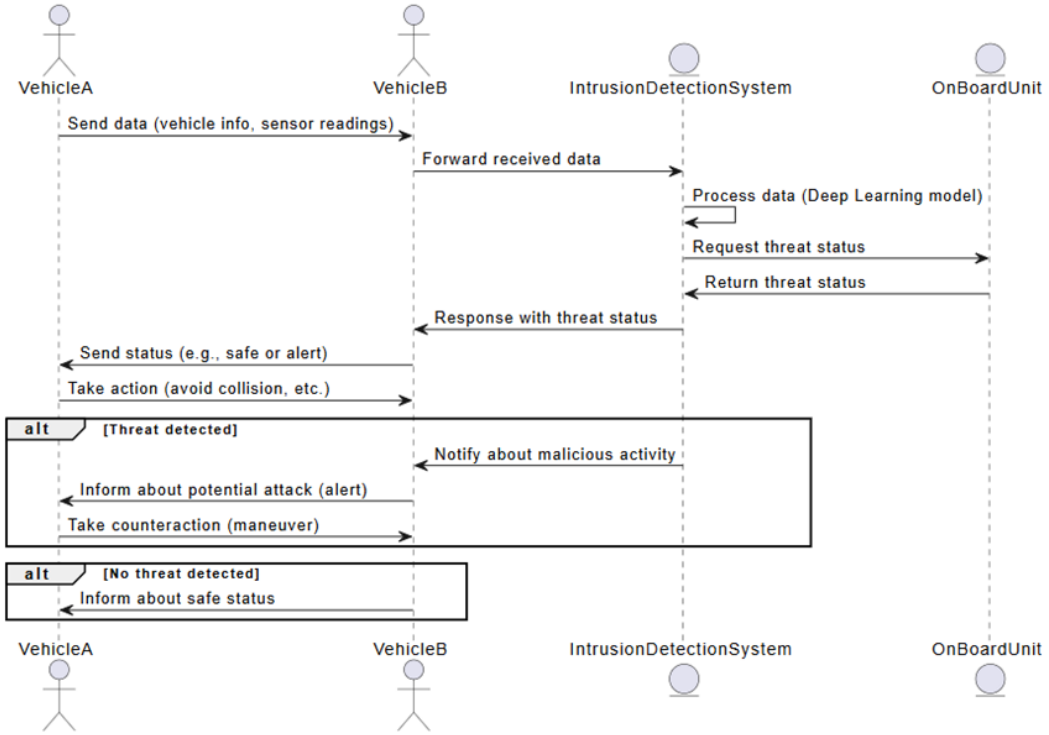
Secondly, A significant limitation of current research is the scarcity of high-quality, labelled intrusion datasets specific to V2V communication. Existing datasets, such as CICIDS2017, primarily address general network security rather than vehicle-specific cyber threats. Future research will explore Generative Adversarial Networks (GANs) to generate synthetic intrusion data, improving model generalization and covering a wider range of potential attack scenarios.

With these advancements, the research can transition from theoretical modelling to practical, real-world applications, enhancing the security of V2X systems.

## 3 Use Case Requirements

- **REQ-01:** It is critical that the V2V communication system is secure from cyber threats such as data injection, jamming, and spoofing, ensuring vehicle safety and preventing accidents caused by malicious intrusions.
- **REQ-02:** It is crucial that the intrusion detection system (IDS) operates in **real-time**, enabling the immediate identification and mitigation of attacks to prevent catastrophic consequences in connected vehicles.
- **REQ-03:** It is essential that the IDS is **scalable** to accommodate large vehicle networks, ensuring that communication between a growing number of vehicles, infrastructure, and pedestrians remains secure without sacrificing performance.
- **REQ-04:** It is critical that the IDS is capable of **generalizing** across various attack scenarios, allowing it to detect both known and unknown threats, thereby ensuring robust security for all potential attack vectors in V2V communication.
- **REQ-05:** It is essential that the IDS can be deployed on **resource-constrained devices**, such as Onboard Units (OBUs), optimizing for computational efficiency while maintaining high accuracy in intrusion detection.

## 4 Sequence Diagram




## 5 References

[1] Khalil, Abizar, Haleem Farman, Moustafa M. Nasralla, Bilal Jan, and Jamil Ahmad. "Artificial Intelligence-Based Intrusion Detection System for V2V Communication in Vehicular Adhoc Networks." Ain Shams Engineering Journal 15, no. 4 (2024): 102616. https://doi.org/10.1016/j.asej.2023.102616.

[2] Rajapaksha, Sampath, Harsha Kalutarage, M. Omar Al-Kadri, Andrei Petrovski, Garikayi Madzudzo, and Madeline Cheah. "AI-Based Intrusion Detection Systems for In-Vehicle Networks: A Survey." ACM Computing Surveys 55, no. 11 (November 2023): Article 237. https://doi.org/10.1145/3570954.

[3] Sowmya, T., and E. A. Mary Anita. "A Comprehensive Review of AI-Based Intrusion Detection System." Measurement: Sensors 28 (2023): 100827. https://doi.org/10.1016/j.measen.2023.100827.

[4] Guerra, Lorenzo, Linhan Xu, Paolo Bellavista, Thomas Chapuis, Guillaume Duc, Pavlo Mozharovskyi, and Van-Tam Nguyen. "AI-Driven Intrusion Detection Systems (IDS) on the ROAD Dataset: A Comparative Analysis for Automotive Controller Area Network (CAN)." In Proceedings of the 2024 Cyber Security in CarS Workshop (CSCS '24), 39–49. New York: Association for Computing Machinery, 2024. https://doi.org/10.1145/3689936.3694696.

[5] AI4CE. V2X-Sim. GitHub. Accessed June 19, 2025. https://github.com/ai4ce/V2X-Sim.

[6] Chethuhn. Network Intrusion Dataset. Kaggle. Accessed June 19, 2025. https://www.kaggle.com/datasets/chethuhn/network-intrusion-dataset.

# Use Case 9: Environment-Based Endurance Testing in Autonomous Vehicles

**Organization**: Sri Sivasubramaniya Nadar College of Engineering

**Country**: India

**Contact Person(s)**:

Hemamala V S - hema2210733@ssn.edu.in

Kiruthick Raja D -kiruthickraja@gmail.com

Kaythry P- kaythryp@ssn.edu.in

## 1 Use Case Summary Table

| Item | Details |
|---|---|
| Category | Intelligent Transport |
| Problem Addressed | Range anxiety and suboptimal route planning in autonomous vehicles due to static battery estimation, lack of terrain-awareness, and absence of real-time intelligence |
| Key Aspects of Solution | • Real-time battery consumption prediction using LSTM / Transformer models.<br>• Context-aware driving assistance<br>• Smart charging station selection using RL. |
| Technology Keywords | Terrain Analysis, Adaptive Routing, Battery Estimation, Deep Learning, Time Series Forecasting, Reinforcement Learning, Real-time Navigation, Smart Charging |
| Data Availability | (Sample dataset for Initial Prototyping phase)<br>• EV battery consumption, Charging Patterns: [8]<br>• Elevation, Congestion, Road Condition: [9]<br>• Temperature, Humidity: [10]<br>• Energy Consumption Benchmark: [11] |
| Metadata Sources and Data collected | Source (Type of Data): Metadata collected<br>Vehicle Sensors (CAN): Speed, SoC, battery temperature, regenerative braking, power draw<br>GPS & Maps API: Coordinates, elevation, road gradient, curvature<br>Traffic APIs: Real-time congestion, average speed, and roadblocks<br>Weather APIs: Temperature, humidity, wind speed, and precipitation<br>Charging Stations: Availability, connector type, estimated queue, energy pricing<br>Driving Behavior: Acceleration/deceleration patterns, idling time, and frequent braking |

(continued)

| Item | Details |
|---|---|
| Model Training and Fine-Tuning | LSTM (Long-Short Term Memory) and Transformer-based models trained on historical driving, road gradient, and battery usage data for energy forecasting; RL models continuously fine-tuned using real-time feedback and usage logs. |
| Testbeds or Pilot Deployments | Not yet deployed; simulated driving environment and historical driving datasets to be used in initial prototyping phase. |
| Code repositories | N/A (To be published upon completion of PoC and initial prototype) |

# 2 Use Case Description

## 2.1 Description

Context & Background:

- **Problem with existing solutions:** The existing solutions/Apps could tell us the route and traffic, and cars shows us the battery level or speed limits – but none of these systems tell us how to drive in advance to maximize our battery, reduce wear, and improve safety based on the upcoming terrain, our battery's health, and our own driving behavior. They're all reactive dashboards – not predictive advisors.
- **What we are doing differently:** Our system takes real-time driver input, vehicle battery condition, and terrain elevation data, and uses AI models (TST + Mistral) to predict: Ideal speed and RPM for the next 5 km, based on battery health, terrain slope, driver behavior, and route curvature. It presents it as a real-time, actionable advisory for the driver.

This isn't just a map. It's a driver- and EV-aware AI assistant that evolves and adapts across driving sessions. Think of it like Cruise Control on Steroids, but made specifically for your country roads, tier-2/tier-3 conditions, and personalized EV optimization.

- **Real-World Example:** Let's say you're entering a steep hill for 2 km. Your battery is overheating, and you tend to brake aggressively. Our system will say: 'Reduce RPM to 1600, hold at 40 km/h for the next 3 km to maintain battery life and avoid overheating. Regenerative braking will kick in on the downhill. No other system tells you this, this early, or this personally.

Objectives & Aim: We aim to develop an **AI-driven real-time endurance estimation framework** in autonomous vehicles.

- To enhance route optimization by combining route, elevation (OpenTopography), and curvature every 5 km with integration of deep learning models with real-time contextual data streams, including road topology, traffic conditions, environmental parameters, and driver-specific behavioural patterns.
- To leverage *time-series forecasting models (LSTM/Transformer)* such as *the Chronus TST model* for time-series prediction of RPM/speed, where the system continuously learns from historical telemetry and energy consumption data to provide accurate, dynamic predictions of battery usage.
- To incorporate Reinforcement Learning (RL) with Mistral LLM (via Ollama) to convert model output into driver instructions to deliver adaptive, personalized recommendations for energy-efficient driving behaviour.

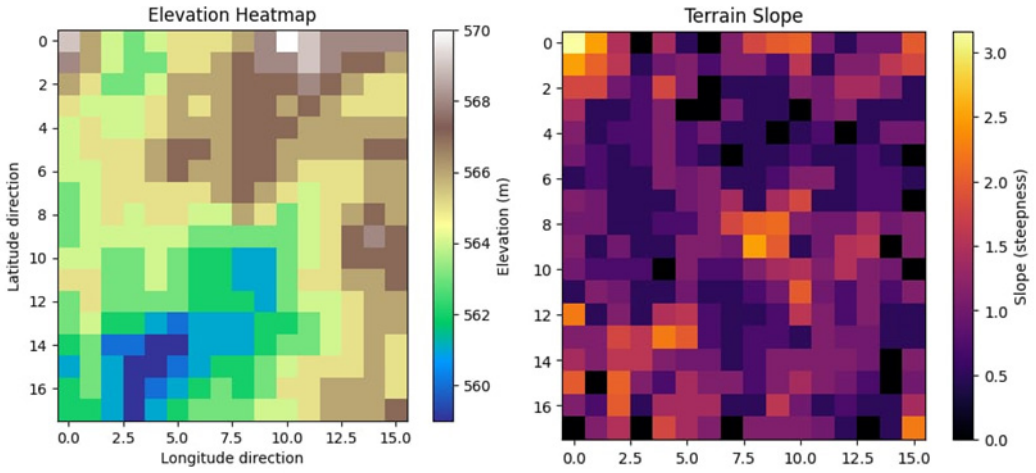
- To make the system continuously learns from historical telemetry and energy consumption data of driver behavior/profiles
- To compute recommendations based on factors such as live battery state (i.e, by reading battery SoC, SoH, temperature, discharge rate – not just SoC, terrain complexity, congestion patterns, and real-time charging slot availability.
- The overarching aim is to create a robust, self-improving system that minimizes energy wastage, reduces range anxiety, and enables autonomous vehicles to make context-aware, data-driven navigation, proactive driving prediction, not just route display and charging decisions in real time.

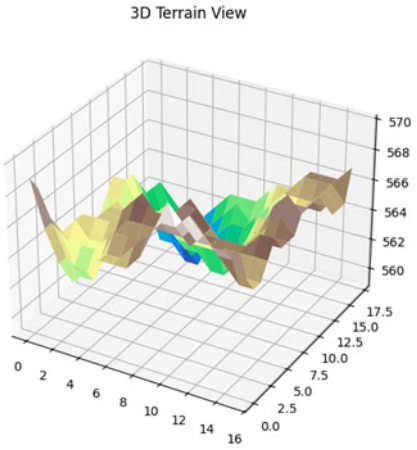
## Sub-Scenarios (Expanded)

To give a more concrete idea of how the system functions in different contexts, we outline the following sub-scenarios:

## Sub-Scenario 1: Energy Optimization

- Uses LSTM/Transformer models to predict high-consumption zones (e.g., steep inclines or traffic-heavy areas).
- Recommends alternate routes or adaptive cruise control strategies to minimize energy drain.



### Sub-Scenario 2: In-Journey Feedback for Driving Style

- Monitors real-time driving behavior (e.g., harsh acceleration or frequent braking).
- Recommends smoother driving patterns using reinforcement learning, which improves range by 8–12%.

Sub-Scenario 3: Smart Journey Planning

- Suggests charging stops based on real-time SoC, terrain, and slot availability.
- Dynamically adjusts the plan if unexpected traffic or battery drain is detected.

Actions Taken (Driving Style Recommendations)

- The system continuously monitors driver input and battery output. Based on this, it suggests:

### Table I – Driving Style Suggestions.

| Situation | Suggested Action |
|---|---|
| High energy usage at an incline | "Reduce speed to 45 km/h" |
| Frequent braking in city driving | "Maintain more distance to avoid stops." |
| Inefficient regen usage | "Lift foot earlier before stop" |
| High temperature impact | "Pre-cool vehicle while plugged in" |

These are communicated via audio (if supported) and visual cues on the HUD.

Innovative Technological approach: The proposed framework introduces a multi-layered AI architecture that integrates deep learning, time-series forecasting, and reinforcement learning to enable real-time, adaptive energy management in autonomous electric vehicles. At the core of this innovation lies a hybrid prediction engine combining LSTM and Transformer models to perform granular battery consumption forecasting by dynamically analysing sequential inputs such as road gradient, elevation, traffic density, environmental conditions, and driver-specific behavioural data. These predictions are further refined through a terrain-aware processing module that translates topographic and environmental metadata into energy expenditure coefficients. The system continuously ingests real-time telematics to fine-tune predictions and employs a reinforcement learning agent to optimize charging decisions by learning from historical feedback, user behaviour, and system performance metrics. This enables the vehicle to recommend context-aware charging stations by evaluating route efficiency, congestion levels, and real-time charger availability. The proposed architecture not only advances battery endurance prediction but also delivers actionable, in-journey driving recommendations through in-vehicle interfaces such as HUDs, creating a closed-loop decision-support system that adapts autonomously to evolving conditions.

Types of Models :

We use a multi-model pipeline, including:

### Table II – Models

| Model type | Usecase |
|---|---|
| Chronos | Transformer , LSTM |
| Mistral | NLP,RL (loss function ) |

Transformer Size & On-Vehicle Deployment

- We use a compressed transformer model—Chronous(T5) with 7B parameters.
- Fine-tuned for battery forecasting and real-time sequence understanding.
- Models are optimized using quantization + pruning.
- In-vehicle edge deployment using a runtime such as TensorRT on an NVIDIA Jetson Nano/Xavier.

Mistral for Natural Language Processing?

Mistral AI has developed models like Mistral 7B, which, despite having fewer parameters than some competitors, deliver performance that rivals larger models. This efficiency is achieved through innovative architectural choices, such as Grouped-Query Attention (GQA) and Sliding Window Attention (SWA), enabling faster inference and reduced memory usage.

Key Advantages:

- Efficiency: Mistral models are designed to be lightweight, making them suitable for deployment in environments with limited computational resources.
- Open-Source Flexibility: Being open-source, Mistral allows for greater customization and transparency, facilitating integration into various systems while maintaining data privacy.
- Competitive Performance: In benchmarks, Mistral 7B has demonstrated performance on par with or exceeding that of larger models like LLaMA 2 13 B.

In Our Application: Mistral's efficiency and open-source nature make it an ideal choice for generating real-time, context-aware driving instructions in our EV assistant system, ensuring quick response times and ease of integration.

Chronos for Time Series Forecasting?

Chronos is a transformer-based framework tailored for time series forecasting. It tokenizes time series data, allowing it to leverage the strengths of language models like T5 for forecasting tasks**.**

Key Advantages:

- Zero-Shot Performance: Chronos has shown strong performance on unseen datasets without the need for extensive retraining, making it adaptable to various forecasting scenarios.
- Long-Term Forecasting: It maintains accuracy over longer prediction horizons, which is crucial for anticipating driving conditions ahead.
- Integration with Transformers: By utilizing transformer architectures, Chronos benefits from advanced sequence modeling capabilities, enhancing its forecasting accuracy.

In Our Application: Chronos enables our system to predict future driving conditions, such as optimal speed and RPM, by analyzing patterns in terrain, battery performance, and driver behavior, thereby enhancing the proactive assistance provided to EV drivers.

Model Selection and System Architecture

Chronos for Energy Forecasting

**Rationale:** Chronos is a transformer-based time series forecasting model that tokenizes time series data, enabling it to leverage the strengths of language models for forecasting tasks. It has demonstrated strong performance on unseen datasets without the need for extensive retraining, making it adaptable to various forecasting scenarios.

Performance Metrics: (Real Mean Square Error & Mean Absolute Error)

- RMSE: 0.001443 [12]
- MAE: 0.001105 [13]

These metrics indicate Chronos's high accuracy in forecasting tasks.[14]

Mistral for Behaviour Adjustment and Decision-Making

**Rationale:** Mistral 7B is a lightweight, open-source language model designed for efficient inference. Its architecture, featuring Grouped-Query Attention and Sliding Window Attention, allows for faster inference and reduced memory usage.

**Integration:** Mistral is utilized to interpret multi-source sensor and prediction data, generating real-time, context-aware driving instructions. This acts as a bridge between predictive AI (Chronos) and human-in-the-loop decision-making.

Model Interaction: Sequential and Hierarchical Integration

Our system employs a hybrid architecture combining sequential and hierarchical model interactions:

1. Sequential Flow:

a. **Input**: Start and end GPS coordinates.[15]

b. **Route Planning:** Mapbox API determines the optimal route.

c. **Terrain Analysis:** The OpenTopography API provides elevation and curvature data.

d. **Energy Forecasting:** Chronos predicts energy consumption and optimal speed/RPM for upcoming segments.

e. **Behavior Adjustment:** Mistral generates natural language instructions based on predictions and driver behavior data.

2. Hierarchical Structure:

a. **Base Layer:** Data acquisition and preprocessing.

b. **Middle Layer**: Predictive modeling (Chronos). [16]

c. **Top Layer:** Decision-making and instruction generation (Mistral).

Terrain-Aware Processing: Translating Topographical Data

**Data Source:** OpenTopography provides high-resolution elevation data. [17]

Processing Method:

- **Rule-Based Functions:** We apply domain-specific rules to translate elevation and curvature data into energy coefficients.
- **GIS Data Integration:** Geospatial data is integrated to enhance the accuracy of terrain analysis.

**Outcome:** This approach allows for accurate estimation of energy requirements based on terrain, improving the reliability of our energy forecasts.

Real-Time Performance on Edge Devices

**Deployment Platforms:** NVIDIA Jetson Nano and Xavier.[18]

Optimization Techniques:

- TensorRT-LLM: Utilized to optimize Mistral's inference performance on NVIDIA hardware. [19]

Performance Metrics:

- Latency: Mistral 7 B-Instruct-v0.1 achieves a latency of approximately 10.2 seconds per inference on Jetson AGX Orin. [20]
- Power Consumption: Jetson AGX Orin operates within a configurable power range of 15W to 50W, suitable for edge deployments.[20]

Evaluation Metrics

Chronos:

- RMSE: 0.001443 [21]
- MAE: 0.001105 [21]

Mistral:

- Latency: Approximately 10.2 seconds per inference on Jetson AGX Orin.[22]

Reinforcement Learning Component:

- Reward Convergence: Monitored through cumulative reward trends over training episodes. [23]

Sensor Inputs for Driving Behavior Analysis

Sensors Utilized:

- Accelerometer: Detects harsh braking and acceleration patterns.
- Gyroscope: Monitors vehicle orientation and stability.
- Battery Management System (BMS): Provides data on battery temperature, state of charge, and health.

Data Processing:

- **Driver-Specific Variations**: Historical driving data is analyzed to personalize recommendations. [24]

**Pattern Recognition:** Machine learning algorithms identify inefficient driving behaviors, enabling targeted feedback.

Impacts & Benefits of the Solutions:

In contrast to existing systems that provide reactive feedback, our solution proactively advises drivers by forecasting vehicle performance and translating these insights into actionable instructions. By leveraging the efficiency of Mistral for language generation and the predictive prowess of Chronos for time series forecasting, we deliver a comprehensive, real-time endurance driver assistance system that enhances safety and efficiency.

**Example:** 'Reduce speed to 45 km/h and maintain RPM around 1900. Estimated range: 8.3 km. The upcoming downhill slope can regenerate ~3.5% battery.

Sample Results (Simulation)

- Route A (no AI): Battery drop of 27% over 62km
- Route B (AI-optimized): Battery drop of 19.8% with 5.6% faster ETA
- Reinforcement-based driving feedback showed ~9.2% improvement in energy efficiency

Visualization:

In the vehicle interface, it (HUD, cluster, center console) shows:

- Predicted Range vs. Actual Range
- Real-time consumption vs. forecast
- Terrain profile and impact overlay
- Suggested action: "Maintain steady acceleration", "Avoid hard braking", etc.

**Impact on Intelligent Transport :**

This project promotes the use of efficient and clean energy by optimizing electric vehicle charging, which helps minimize energy waste. It leverages advanced AI-driven infrastructure to support and accelerate the adoption of electric vehicles, fostering innovation within the transportation sector.

By improving charging processes, the project helps reduce urban congestion and lowers harmful emissions, contributing to more sustainable and livable cities. Additionally, the focus on reducing energy waste directly supports broader climate action efforts, enhancing sustainability and resilience.

Overall, this initiative aligns with global goals for affordable clean energy, innovative infrastructure, sustainable urban development, and climate protection.

**Partner name: Yes, Potential Partners: AV manufacturers, smart grid providers.**

## 2.2 Benefits of use case

In contrast to existing systems that provide reactive feedback, our solution proactively advises drivers by forecasting vehicle performance and translating these insights into actionable instructions. By leveraging the efficiency of Mistral for language generation and the predictive prowess of Chronos for time series forecasting, we deliver a comprehensive, real-time endurance driver assistance system that enhances safety and efficiency.

**Example:** 'Reduce speed to 45 km/h and maintain RPM around 1900. Estimated range: 8.3 km. The upcoming downhill slope can regenerate ~3.5% battery.

Sample Results (Simulation)

- Route A (no AI): Battery drop of 27% over 62km
- Route B (AI-optimized): Battery drop of 19.8% with 5.6% faster ETA
- Reinforcement-based driving feedback showed ~9.2% improvement in energy efficiency

Visualization:

In the vehicle interface, it (HUD, cluster, center console) shows:

- Predicted Range vs. Actual Range
- Real-time consumption vs. forecast
- Terrain profile and impact overlay
- Suggested action: "Maintain steady acceleration", "Avoid hard braking", etc.

This project promotes the use of efficient and clean energy by optimizing electric vehicle charging, which helps minimize energy waste. It leverages advanced AI-driven infrastructure to support and accelerate the adoption of electric vehicles, fostering innovation within the transportation sector.

By improving charging processes, the project helps reduce urban congestion and lowers harmful emissions, contributing to more sustainable and livable cities. Additionally, the focus on

reducing energy waste directly supports broader climate action efforts, enhancing sustainability and resilience.

Overall, this initiative aligns with global goals for affordable clean energy, innovative infrastructure, sustainable urban development, and climate protection.

### 2.3 Future work

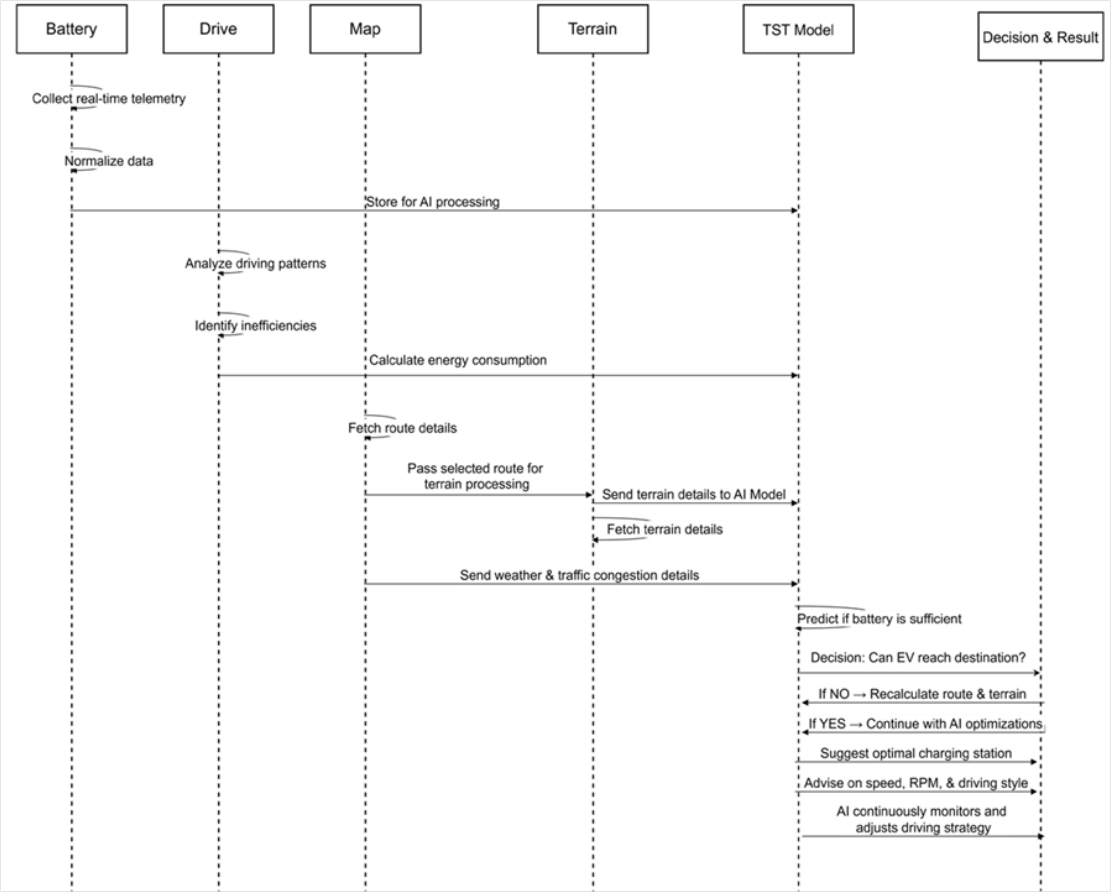
With funding, we will proceed with:

- Extensive Data Collection: Real-world driving data from EV fleets to improve model accuracy.
- Advanced ML Techniques Integrating Graph Neural Networks (GNNs) to improve terrain-aware battery predictions.
- Deployment & Real-Time Testing: Deploying AI models in edge devices inside EVs to provide instant predictions.
- Collaboration with Charging Networks, Integrating AI-driven booking & reservation systems for charging slots

## 3 Use Case Requirements

- **REQ-01**: *It is mandatory* that the system continuously acquires, processes, and integrates real-time data from vehicle telemetry, GPS, traffic APIs, road gradient, environmental sensors, and charging infrastructure through standardized APIs or onboard modules for context-aware decision-making.
- **REQ-02**: *It is critical* to implement an AI-based terrain-aware energy forecasting model (Chronos) that predicts and analyzes energy consumption using historical data, real-time telemetry, terrain profiles, and driver behavior patterns.
- **REQ-03**: *It is critical* that the reinforcement learning-driven optimization engine provides adaptive, personalized driving and charging recommendations by learning continuously from live driving conditions, battery status, environmental data, and historical feedback.
- **REQ-04**: *It is mandatory* that the system supports an edge-optimized, low-latency deployment architecture with in-vehicle edge devices capable of sub-second latency for data ingestion, prediction, and feedback, incorporating fallback lightweight models for offline operability where the Mistral (RL-based NLP model) calculates the loss function.
- **REQ-05**: *It is critical* to integrate a context-aware driver assistance and visualization interface using the HUD, cluster, and center console to deliver actionable recommendations and visualize live and predicted battery status, range, route, and upcoming terrain changes.
- **REQ-06**: *It is mandatory* that resilient, redundant, and highly available APIs are used for route, traffic, weather, and charging data, ensuring secure, standardized, and anonymized data handling aligned with data privacy guidelines, even under connectivity loss.
- **REQ-07**: *It is critical* to include continuous model evaluation and a scalable simulation environment with a retraining pipeline and loss function monitoring to detect model drift and validate AI model behavior.

## 4 Sequence Diagram




## 5 References

[1] ITU. *AI for Good–Innovate for Impact Report 2024*, Use Case 36: "Bidirectional Telecom-to-Grid and Telecom-to-Vehicle Energy Exchange for Net-Zero." Accessed June 19, 2025. https://demo.ifgict.org/wp-content/uploads/2024/09/ITU-AI-for-Good-Innovate-final-report-1.pdf.

[2] Zhu, Z., S. Gupta, A. Gupta, and M. Canova. "A Deep Reinforcement Learning Framework for Eco-Driving in Connected and Automated Hybrid Electric Vehicles." *IEEE Transactions on Vehicular Technology* 73, no. 2 (February 2024): 1713–1725. https://doi.org/10.1109/TVT.2023.3318552.

[3] Comma.ai. "Openpilot: An Open-Source Driving Automation System." Accessed June 19, 2025. https://comma.ai/openpilot.

[4] Kerbel, L., B. Ayalew, A. Ivanco, and K. Loiselle. "Driver Assistance Eco-driving and Transmission Control with Deep Reinforcement Learning." In *2022 American Control Conference (ACC)*, 2409–2415. Atlanta, GA, USA, 2022. https://doi.org/10.23919/ACC53348.2022.9867557.

[5] Holden, J., N. Reinicke, and J. Cappellucci. "Route E: A Vehicle Energy Consumption Prediction Engine." *SAE International Journal of Advanced & Current Practices in Mobility* 2, no. 5 (2020): 2760–2767. https://doi.org/10.4271/2020-01-0939.

[6] Lombardi, Federico, et al. "RAMP: Stochastic Simulation of User-Driven Energy Demand Time Series." *Journal of Open Source Software* 9, no. 98 (2024). https://doi.org/10.21105/joss.06418.

[7] Yong, P., et al. "Evaluating the Dispatchable Capacity of Base Station Backup Batteries in Distribution Networks." *IEEE Transactions on Smart Grid* 12, no. 5 (September 2021): 3966–3979. https://doi.org/10.1109/TSG.2021.307475.

[8] National Renewable Energy Laboratory (NREL). *EV Battery Consumption Charging Patterns*. Accessed June 19, 2025. https://www.nrel.gov/docs/fy24osti/85902.pdf.

[9] Google Earth Engine. *Elevation, Congestion, Road Condition Datasets*. Accessed June 19, 2025. https://developers.google.com/earth-engine/datasets.

[10] NASA POWER. *Temperature and Humidity Data Access Viewer*. Accessed June 19, 2025. https://power.larc.nasa.gov/data-access-viewer/.

[11] Bureau of Energy Efficiency (India). *Energy Consumption Benchmark Reports*. Accessed June 19, 2025. https://beeindia.gov.in/en/publications#id-reports-studies.

[12] GitHub. *Chronos Forecasting RMSE*: *0.001443*. Accessed June 19, 2025. https://github.com/.

[13] FG Research. *Chronos Forecasting MAE: 0.001105*. Accessed June 19, 2025.

[14] arXiv. *Chronos Accuracy Metrics (RMSE, MAE)*. Accessed June 19, 2025. https://arxiv.org/.

[15] OpenTopography. *Sequential Flow Input: Start and End GPS Coordinates*. Accessed June 19, 2025. https://opentopography.org/.

[16] arXiv. Hierarchical Architecture: *Middle Layer–Predictive Modelling*. Accessed June 19, 2025. https://arxiv.org/.

[17] OpenTopography. *Terrain-Aware Processing Data Sources*. Accessed June 19, 2025. https://opentopography.org/.

[18] Mistral Solutions. *Deployment Platforms: NVIDIA Jetson Nano and Xavier*. Accessed June 19, 2025.

[19] Mistral Solutions. *Tensor LLTM Performance Optimization*. Accessed June 19, 2025. https://github.com/.

[20] Mistral Solutions. *Edge Device Performance Metrics: Latency and Power Consumption*. Accessed June 19, 2025. https://developer.nvidia.com/.

[21] FG Research. *Chronos Evaluation Metrics: RMSE and MAE*. Accessed June 19, 2025.

[22] Mistral Solutions. *Mistral Evaluation Metrics: Latency*. Accessed June 19, 2025. https://developer.nvidia.com/.

[23] NVIDIA. *Deep Reinforcement Learning in Robotics with Jetson: Reward Convergence*. Accessed June 19, 2025.

[24] FG Research. *Driver-Specific Variations in Chronos Forecasting*. Accessed June 19, 2025.

# 4.11. Category 11: Smart Agriculture

## Use Case 1: AI-Enabled Soil Analysis and Weather Station for Local Farmers

**Organization:** CSIR - Institute for Scientific and Technological Information

**Country**: Ghana

**Contact Person(s)**:

Dennis Agyemanh Nana Gookyi, dennisgookyi@gmail.com,

Fortunatus Aabangbio Wulnye,

Michael Wilson,

John Awotwi,

Yaw Twum Barimah,

Paul Danquah,

Roger Kwao Ahiadormey

### 1 Use Case Summary Table

| Item | Details |
|---|---|
| Category | Smart Agriculture |
| Problem Addressed | Smallholder farmers often lack access to real-time, localized soil and weather data, which hinders their ability to make informed decisions about crop management. Traditional data collection methods are typically slow, imprecise, or difficult to access, resulting in inefficient farming practices, suboptimal resource use, and reduced agricultural productivity. This project seeks to address this critical gap by delivering accurate, real-time soil and weather insights to support data-driven decision-making for both farmers and agricultural policymakers. |
| Key Solution | Real-time soil and weather data including temperature, humidity, pH, NPK, and rainfall are collected using fixed sensors without moving parts. This data powers AI models that generate localized forecasts, while a TinyML-enabled device provides on-site crop recommendations based on current soil conditions. All data is visualized through Grafina dashboards and made publicly accessible to support informed decision-making by farmers, researchers, and policymakers. |

(continued)

| Item | Details |
|---|---|
| Technology Keywords | Artificial Intelligence (AI), soil analysis, weather station, real-time data, predictive modelling, IoT, sensors |
| Data availability | Public - The data will be available upon completion of the project |
| Metadata (type of data) | Environmental data (rainfall, temperature, humidity, pressure)<br>Soil data (temperature, humidity, pH, NPK, conductivity) |
| Model Training and fine tuning | Supervised machine learning models are trained on soil and weather data to enable accurate predictions. As data increases, lightweight deep learning models are fine-tuned for TinyML deployment, allowing real-time, low-power crop recommendations directly from the sensor station without constant cloud connectivity. |
| Testbeds or Pilot Deployments | Pilot tested by Team GenStorm using Arduino-based TinyML models for smart weather monitoring, as documented [1] |
| Code repositories | Code repository containing Arduino implementation, TinyML model files, dataset, and documented [2] |

# 2 Use Case Description

## 2.1 Description

Access to real-time, localized soil and weather data is a major challenge for farmers, especially in rural and underserved regions. Without accurate and timely information, farmers struggle to make informed decisions about planting, fertilization, irrigation, and harvesting, leading to lower yields, resource inefficiencies, and environmental degradation. Traditional soil testing methods are slow, expensive, and often inaccessible, while available weather data typically lacks the local precision needed for effective farm management.

The main objective of this project is to bridge this gap by developing an AI-enabled soil analysis and weather station that delivers real-time, actionable data to farmers. Specifically, the project aims to (1) collect live soil and weather metrics, (2) train predictive models tailored to local farm conditions, (3) deploy a TinyML-enabled device offering real-time, data-driven crop recommendations, and (4) make the generated data publicly available to inform policymaking and support evidence-based agricultural strategies.

The technological approach combines Internet of Things (IoT) sensor networks, AI-based predictive modeling, and edge computing using TinyML. The soil and weather station, designed without moving parts for durability, collects critical parameters such as rainfall intensity, soil pH, temperature, and NPK levels. Data will be streamed in real-time to cloud servers using AWS, visualized through the Grafina platform, and used to train machine learning models that are later fine-tuned for lightweight, on-device inference. By using TinyML, the system minimizes dependency on constant cloud access, ensuring farmers receive instant recommendations even in remote locations.

The expected impact includes enhanced agricultural productivity through data-driven decision-making, optimized use of fertilizers and water, improved crop yields, and more sustainable

farming practices. On a broader scale, publicly available data will empower policymakers to implement targeted interventions, promote farmland restoration, and drive innovation in post-harvest technologies, ultimately contributing to food security and environmental sustainability in Ghana and beyond.

Use Case Status: Ongoing

## 2.2 Benefits of the use case

Improve the livelihoods of local farmers, who often live in poverty and rely solely on crop sales for income. By providing AI-enabled predictive soil and weather insights, farmers can better plan their planting and cultivation cycles, leading to increased yields and improved economic outcomes.

Support climate action through the collection and analysis of localized environmental data from farms across the country. This data helps visualize the impact of climate change on agriculture and informs targeted interventions to mitigate its effects.

## 2.3 Future Work

Several enhancements and developments are planned to advance the project beyond its current experimental phase. The immediate next steps include the integration of GPS functionality to enable precise location tagging of collected data, improving the spatial accuracy of soil and weather monitoring. Additional field deployments across multiple farms and diverse geographical regions are planned to expand the data set and enhance model generalization.

Further work will focus on upgrading sensor hardware for better durability and precision, particularly for long-term outdoor deployments. The machine learning models will undergo continuous improvement through expanded data collection, feature engineering, and the exploration of advanced algorithms such as ensemble learning and deep neural networks optimized for TinyML deployment.

Additional resources needed include advanced IoT modules, low-power edge AI chips, extended AWS cloud services for scalable data management, and increased access to diverse soil samples from across different agro-ecological zones. Collaboration with local agricultural extension services, universities, and international research institutions is envisioned to support model validation, farmer engagement, and technology transfer.

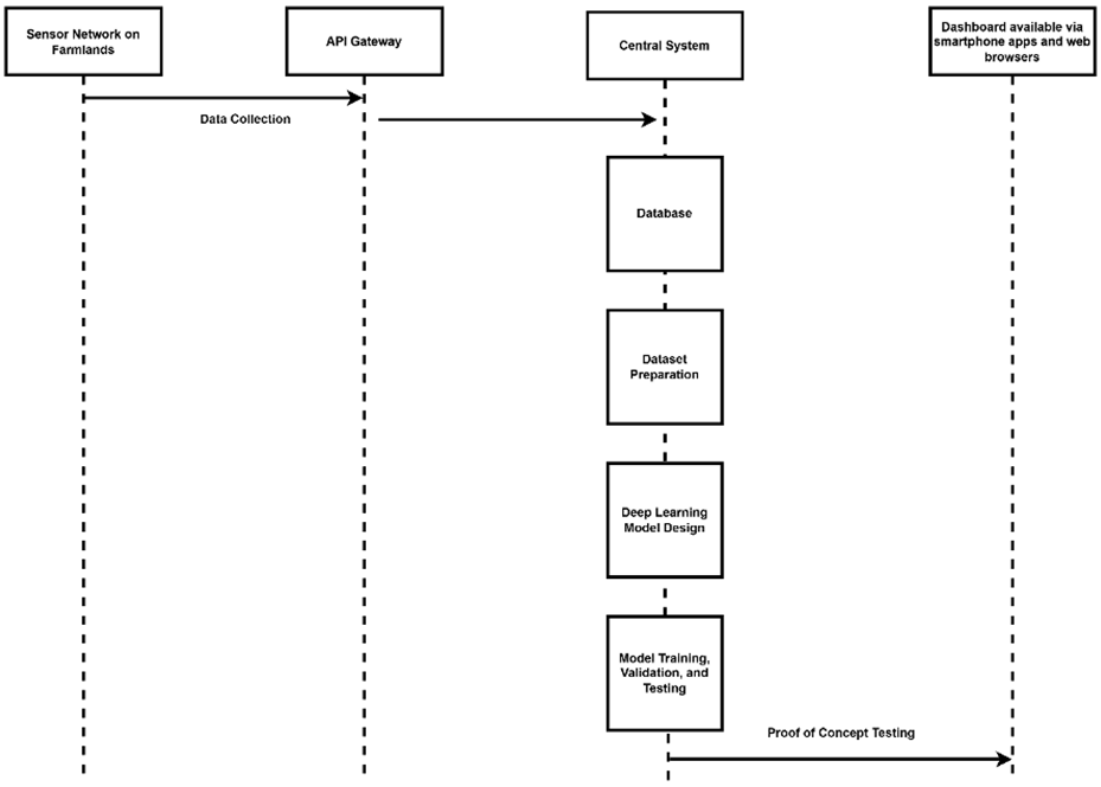
Future expansions will also explore integrating pest and disease monitoring into the station's capabilities, enabling even broader support for smart farming practices. By fostering open collaborations and sharing findings, the project aims to create a scalable model for data-driven agriculture that can be replicated in other regions facing similar challenges.

# 3 Use Case Requirements

- **REQ - 01:** It is required to collect real-time soil metrics (temperature, humidity, pH, NPK, and conductivity) and weather metrics (rainfall, temperature, humidity, and pressure) using fixed sensors without moving parts.
- **REQ - 02:** It is required to securely and reliably transmit the collected data to the cloud for storage and analysis.

- **REQ - 03:** It is required to integrate with the Grafina platform to enable real-time data visualization and dashboard generation.
- **REQ - 04:** It is required to implement server-side timestamping for accurate and traceable data logging.
- **REQ - 05:** It is required to develop supervised machine learning models trained on soil and weather data to enable localized predictions.
- **REQ - 06:** It is required to optimize and deploy lightweight TinyML models for low-power, on-device crop recommendation delivery.
- **REQ - 07:** It is recommended to validate and calibrate soil sensors to ensure consistent accuracy across various soil types.
- **REQ - 08:** It is recommended to provide public access to anonymized, aggregated data to support research, policymaking, and community initiatives.
- **REQ - 09:** It is recommended to include GPS functionality to enable location-specific data tagging and services.
- **REQ - 10:** It is recommended to use cloud infrastructure for scalable computing, machine learning and model training.

## 4 Sequence Diagram



## 5 References

[1] IEEE. "Power Efficient Non-Mechanical Weather Station: Harnessing TinyML for Rainfall Classification and Environmental Sensing." *IEEE Xplore*. https://ieeexplore.ieee.org/document/10856468.

[2] International Telecommunication Union (ITU). *Next-Gen TinyML Smart Weather Station Challenge 2024*. Accessed June 18, 2025. https://challenge.aiforgood.itu.int/match/matchitem/91.

# Use Case 2: IoT-Based Solar-Powered Automatic Basin Irrigation System

**Organisation:** Water Technology Centre, ICAR - Indian Agricultural Research Institute

**Country**: India

**Contact Person(s)**:

Dr Monalisha Pramanik - monalishapramanik@gmail.com

Dr Manoj Khanna - khanna_manoj2001@yahoo.com.au

Dr P S Brahmanand - psba12345@gmail.com

## 1 Use Case Summary Table

| Item | Details |
|---|---|
| Category | Smart Agriculture |
| Problems Addressed | Traditional basin irrigation systems face several challenges that hinder their efficiency and effectiveness. One major issue is poor water application efficiency, often caused by deep percolation losses. These systems also require labor-intensive efforts to monitor and control water flow during irrigation events, making them time-consuming and physically demanding. Additionally, uneven water distribution frequently results in under-irrigation or over-irrigation of crops, negatively affecting plant health and yields. Collectively, these issues lead to the wasteful use of essential farming resources such as water, energy, and time. |
| Key Solution | Automatic Basin Irrigation System addresses the limitations of traditional methods by incorporating real-time field monitoring through wireless Internet of Things (IoT) connectivity using LoRa and GSM networks, alongside soil moisture sensors. Irrigation is automatically scheduled and controlled based on preset soil moisture thresholds, ensuring timely and efficient water delivery. Microcontrollers, coupled with mobile and web dashboards, enable accurate control and monitoring of water supply. This data-driven approach significantly reduces labor demands while optimizing water usage. As a result, the system enhances irrigation application efficiency and improves distribution uniformity, leading to more sustainable and productive farming practices. |
| Technology Keywords | Wireless, Soil moisture Sensor, Internet of Things, Automation, Irrigation |

(continued)

| Item | Details |
|---|---|
| Data Availability | Not publicly available. |
| Metadata (type of data) | Date and Time, Soil moisture sensor ID, Location (Lat long), Unit of Measurement (%), Soil moisture data (JSON and CSV format ), Device Status (Online, Battery), Gate Open/Close |
| Model Training and Fine-Tuning | Collected soil moisture, device status, and environmental data offer strong potential for developing predictive models to optimize irrigation schedules, detect anomalies, and forecast soil moisture trends. |
| Testbeds or Pilot Deployments | The system was tested and deployed at the research farm, IARI, New Delhi. |

# 2 Use Case Description

## 2.1 Description

Surface irrigation remains the most widely used method for watering crops and pastures both in India and globally, largely due to its low cost and minimal energy requirements. In India, this method plays a crucial role in ensuring the availability of food and fodder. For example, during 2016 - 2017, approximately 80% of the country's total irrigated land was managed through surface irrigation systems.

Despite its widespread use, traditional surface irrigation systems are often associated with low water application efficiency and high labor demands. Although well designed and managed systems have the potential to achieve high efficiency, many existing systems continue to operate at low and highly variable efficiency levels. Given its cost-effectiveness and prevalence, surface irrigation is expected to remain a dominant irrigation method both in India and internationally for the foreseeable future. Consequently, improving the application efficiency of these systems remains a critical area of focus.

Automation using IoT technologies presents a promising solution for enhancing the performance of conventional surface irrigation systems. An automated surface irrigation system operates with minimal manual input using timers, sensors, computers, or mechanical devices. Research has shown that incorporating intelligent irrigation controllers and wireless sensor networks can reduce water usage by up to 38% in irrigation projects.

IoT-based systems empower farmers to overcome challenges related to monitoring and labor shortages. Innovative and cost-effective IoT applications are increasingly being adopted in agriculture to improve quality, productivity, sustainability, and cost efficiency. IoT refers to a network of physical devices embedded with electronics, software, sensors, and network connectivity that enables them to collect and exchange data. In such systems, a central unit (sink) interprets the data and sends commands to actuators to execute tasks.

Several studies have demonstrated that IoT systems equipped with wireless sensor networks (WSNs) using communication protocols such as ZigBee or LoRa can effectively transmit field data such as soil moisture and temperature at regular intervals. These systems support irrigation scheduling through mobile applications for various crops.

To address existing research gaps and enhance the efficiency of basin irrigation systems through automation, the following objectives were set:

- To develop a sensor-based automatic basin irrigation system.
- To identify the optimal placement of soil moisture sensors in the field.

### Benefits of the Use Case

Enhance food security by improving irrigation efficiency through an IoT-based, solar-powered basin irrigation system. By delivering water precisely when and where it is needed, the system increases crop productivity and supports sustainable agricultural practices.

Promote sustainable energy use by leveraging low-cost solar power to run the irrigation system, reducing dependency on unreliable or expensive energy sources and enabling continuous operation in remote farming areas.

## 2.2 Benefits of the use case

Enhance food security by improving irrigation efficiency through an IoT-based, solar-powered basin irrigation system. By delivering water precisely when and where it is needed, the system increases crop productivity and supports sustainable agricultural practices.

Promote sustainable energy use by leveraging low-cost solar power to run the irrigation system, reducing dependency on unreliable or expensive energy sources and enabling continuous operation in remote farming areas.

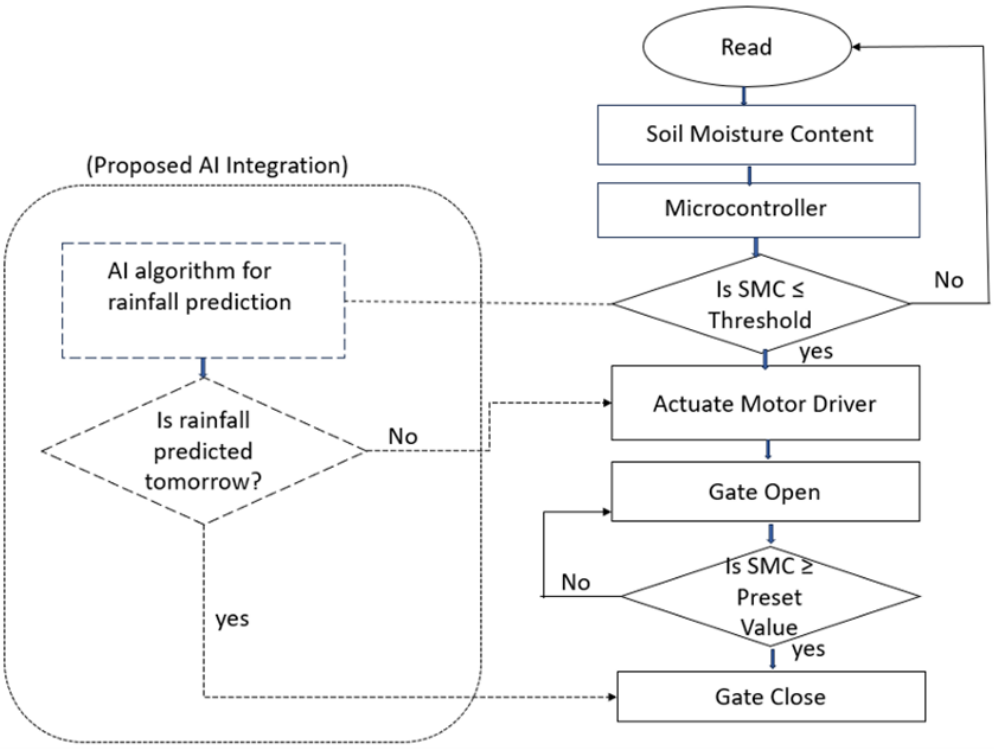
## 2.3 Future Work

The automatic basin irrigation system will be enhanced by integrating weather forecast APIs to utilize real-time data such as rainfall predictions and temperature for smarter irrigation decisions, avoiding unnecessary watering when rain is anticipated. Incorporating nutrient sensors will further support the maintenance of soil health. Pilot deployments across different crops and climatic zones are planned to evaluate the system's scalability, cost-effectiveness, and potential for wider adoption.

# 3 Use Case Requirements

- **REQ-01:** It is required to deploy wireless soil moisture sensors for continuous and accurate monitoring of soil conditions in the field.
- **REQ-02:** It is required to install a check gate mechanism to control and regulate water flow within the irrigation channel.
- **REQ-03:** It is required to use a gateway device to collect sensor data and transmit it securely to the cloud server for storage and analysis.
- **REQ-04:** It is required to develop a mobile based application that allows users to access real-time soil moisture data and schedule irrigation events remotely.

## 4 Sequence Diagram




## 5 References

[1] Monalisha Pramanik, Manoj Khanna, Man Singh, D.K. Singh, Susama Sudhishri, Arti Bhatia, Rajeev Ranjan,(2022) Automation of soil moisture sensor-based basin irrigation system, Smart Agricultural Technology,Volume 2, 2022, 100032, ISSN 2772-3755, https://doi.org/10.1016/j.atech.2021.100032.

[2] Monalisha Pramanik, Manoj Khanna, Man Singh, D K Singh, Susama Sudhisri, Arti Bhatia and Rajeev Ranjan (2022) Performance of sensor-based automatic basin irrigation system in wheat crop under semi-arid conditions. Journal of Agricultural Physics, Vol.2,199-207.

# Use Case 3: AI-Powered Agricultural Contact Center for Farmers

**Organization:** Postal & Telecommunications Regulatory Authority of Zimbabwe (POTRAZ)

**Country:** Zimbabwe

**Contact Person**:

Frank Makeba frankmakeba0@gmail.com, fmakeba@cut.ac.zw

## 1 Use Case Summary Table

| Item | Details |
|---|---|
| Category | Smart Agriculture |
| Problem Addressed | Farmers in Zimbabwe face limited access to timely, localized agricultural advice due to a critically low extension officer-to-farmer ratio of 1:2000 well below the Food Agriculture Organization (FAO) recommended 1:400. Combined with climate-related shocks and a 30% post harvest loss rate, this lack of support severely impacts productivity. While over 63% of the population depends on agriculture, 76% of smallholder farmers live below the poverty line, and 31% of rural adults are illiterate making traditional digital tools ineffective without voice support. These challenges fuel food insecurity, poverty, and environmental degradation, threatening national progress toward key Sustainable Development Goals (SDG) targets. |
| Key Aspects of Solution | An Artificial Intelligence (AI) powered, voice assisted contact center designed to empower farmers with real time agricultural solutions. Farmers can log their problems through phone calls, SMS, or a website, and receive automated voice responses with actionable solutions tailored to their specific needs, local language support, compatibility with 2G devices, inclusive access via SMS/voice for illiterate users. |
| Technology Keywords | Natural Language Processing (NLP), Speech-to-Text, Voice AI, Local Language AI, Predictive Analytics, Chatbot, Climate AI, Internet of Things (IoT) - Ready |
| Data Availability | Data will be available upon request - Agricultural extension datasets from government and Non Government Organizations (NGO), Localized language data, Climate Application Program Interfaces (OpenWeatherMap), Crop disease |
| Metadata (Type of Data) | Audio (calls), Text (SMS/web), Visual (satellite & weather data), Structured (market prices) |

(continued)

| Item | Details |
|---|---|
| Model Training and Fine-Tuning | Fine-tuned Transformer models (NLP & Text To Speech) on local language corpora; GPT-2 custom-trained on Shona data (language in Zimbabwe), disease prediction model trained using Convolutional Neural Networks (CNN) + local datasets. |
| Testbeds or Pilot Deployments | 3 pilot wards in Mashonaland West and Masvingo provinces in Zimbabwe; Evaluation in partnership with Chinhoyi University Agro Industrial Park. |
| Code repositories | ThulaniTheProgrammer, *Hurudzaii*, GitHub [1] |

# 2 Use Case Description

## 2.1 Description

The AI-Powered Agricultural Contact Center for Farmers was developed in response to the deepening crisis in African agriculture, where farmers who form the backbone of the continent's food systems lack timely, localized, and affordable access to expert farming advice, with a collapsing agricultural extension network, rising climate variability, and low levels of digital literacy, most farmers rely on outdated practices or word-of-mouth guidance. AI-Powered Agricultural Contact Center aims to democratize agricultural knowledge by providing a multilingual, AI-powered voice platform that delivers real-time, expert support through voice calls and SMS, accessible even on basic 2G-enabled phones. By supporting indigenous languages like Shona and Ndebele (languages in Zimbabwe), the solution ensures inclusivity and relevance in rural contexts. Technologically, the platform integrates fine-tuned local language models based on GPT-2, text-to-speech systems, and APIs for weather, pest, and market data. It is designed for scalability and optimized for offline-first environments, enabling AI-driven crop diagnostics, fertilizer recommendations, and market insights.

Our AI-Powered Agricultural Contact Center solution has a greater global impact standard. It introduces the first AI-powered multilingual voice interface for farmers in Zimbabwe , projects a 40% reduction in post-harvest losses and a 20% increase in eco-friendly farming practices and targets the digital empowerment of over 10,000 rural farmers by 2026 70% of whom will access the service on low-end phones. By addressing SDGs 1, 2, 9, and 13, the platform is expected to raise household incomes by 25%, improve food security, and strengthen resilience to climate shocks . AI-Powered Agricultural Contact Center also redefines the agricultural extension model by bridging the 1:2000 extension officer gap with always available, AI driven support while breaking access barriers for illiterate, elderly, and disabled users through voice-based interaction in native languages. Our mission is reaching 50% female participation by year three and scaling across the Southern African Development Community (SADC) region through open-source NLP toolkits, our AI-Powered Agricultural Contact Center stands as a transformative and inclusive digital innovation for social good.

**Use Case Status:**

The use case is part of a large product development

## 2.2 Benefits of the use case

Strengthen rural livelihoods by increasing agricultural productivity, reducing input costs through AI-generated recommendations, and improving access to market information. These combined efforts help smallholder farmers boost incomes and achieve greater financial stability.

Enhance food security by providing tailored crop management advice, early disease prediction, and support for climate-resilient practices. This ensures steady food production and reduces hunger among vulnerable farming communities.

Promote gender inclusion by empowering women farmers with accessible voice-based agricultural services that overcome literacy and digital barriers. The project also commits to increasing female participation in training and deployment phases.

Drive digital innovation in underserved regions by delivering intelligent agricultural support via low-tech devices and 2G networks, demonstrating scalable infrastructure solutions for inclusive rural development.

**Partners:** Chinhoyi University of Technology, Postal & Telecommunications Regulatory Authority of Zimbabwe (POTRAZ), Ministry of Agriculture (Zimbabwe), Zimbabwe Centre for High Performance Computing

## 2.3 Future work

Future development will focus on transitioning the platform to a microservices-based, cloud-native architecture to support seamless localization, scalability, and deployment across different countries and languages. A key enhancement will be the integration of code-switching capabilities to reflect how farmers naturally alternate between local languages and regional dialects, ensuring more relatable and effective communication. Strategic partnerships will be explored with organizations such as FAO, World Food Program (WFP), International Fund for Agricultural Development (IFAD), United States Agency for International Development (USAID), and Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) to strengthen farmer support services, access funding opportunities, and ensure long-term sustainability.

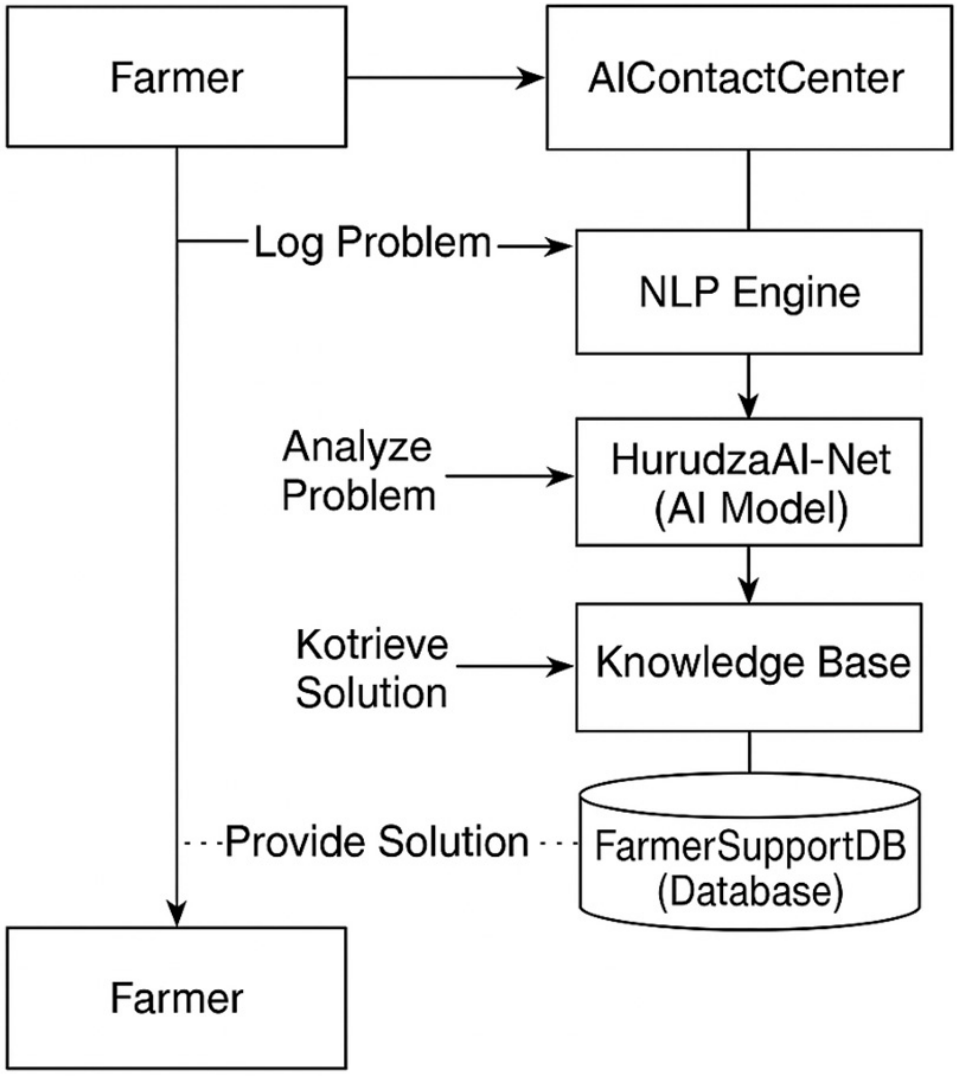
Between 2026 and 2027, the platform will target initial expansion across Africa, beginning with Southern Africa (Zambia, Malawi, Mozambique, Botswana), East Africa (Kenya, Tanzania, Uganda, Rwanda), West Africa (Nigeria, Ghana, Côte d'Ivoire, Senegal), North Africa (Egypt, Morocco), and the Horn of Africa (Ethiopia, Somalia). Looking beyond the continent, low-resource countries such as Haiti, Nepal, Laos, and Timor-Leste are also identified as potential beneficiaries for future deployment of this inclusive agricultural technology.

# 3 Use Case Requirements

- **REQ-01:** It is required to deploy a scalable AI powered chatbot with voice interface optimized for low bandwidth environments to ensure accessibility across rural areas.
- **REQ-02:** It is required to integrate with mobile network APIs to enable USSD and SMS based interactions, supporting farmers with basic feature phones.
- **REQ-03:** It is required to implement disease prediction capabilities using timeseries image and sensor data for early intervention and crop protection.

- **REQ-04:** It is required to provide real time weather alerts to users via SMS and voice channels to support timely decision making.
- **REQ-05:** It is required to develop a multilingual natural language processing engine fine tuned on indigenous languages such as Shona and Ndebele for inclusive communication. **REQ-06:** It is recommended to establish a market data ingestion system linked to agricultural marketplaces to offer price trends and marketing insights.
- **REQ-07:** It is recommended to include an offline fallback mechanism with scheduled call back responses to ensure continuous support in areas with unstable connectivity.

## 4 Sequence Diagram




## 5 References

[1] CCARDESA. *Digital Agriculture Country Study Annex: Zimbabwe*. Gaborone: Centre for Coordination of Agricultural Research and Development for Southern Africa (CCARDESA), 2022. PDF.

[2] Sunday News. "Zimbabwe Ranked High in Digital Agriculture Innovation." *Sunday News (Zimbabwe)*, 2023. https://www.sundaynews.co.zw/zimbabwe-ranked-high-in-digital-agriculture-innovation/.

[3] The Herald. "Digitalisation of Agriculture Sector Critical for Sustainable Production." *The Herald (Zimbabwe)*, 2023. https://www.herald.co.zw/digitalisation-of-agriculture-sector-critical-for-sustainable-production/.

[4] USAID/Strategic Impact Advisors. *Zimbabwe Digital Agriculture Assessment: A Report for USAID/Zimbabwe*. 2023. PDF.

[5] CCARDESA. "South Africa, Zimbabwe and Tanzania Rank High in Digital Agriculture Innovation in the SADC Region." CCARDESA, 2022. https://www.ccardesa.org/news/south-africa-zimbabwe-and-tanzania-rank-high-digital-agriculture-innovation-sadc-region.

[6] CCARDESA. "Digital Agriculture Innovations." *CCARDESA*, 2023. https://www.ccardesa.org/digital-agriculture-innovations.

[7] The Herald. "Watch: Youth Leverage AI in Agriculture Innovation." *The Herald (Zimbabwe)*, n.d. https://www.heraldonline.co.zw/watch-youth-leverage-ai-in-agriculture-innovation/.

[8] CITE. "Youths Urged to Tackle Climate Change with Innovative Ideas." *Center for Innovation and Technology*, n.d. https://cite.org.zw/youths-urged-to-tackle-climate-change-with-innovative-ideas/.

[9] ThulaniTheProgrammer. *Hurudzaii. GitHub*. Accessed June 19, 2025. https://github.com/ThulaniTheProgrammer/Hurudzaii.git.

# Use Case 4: Divine Farmer Agricultural AI Model

**Organization:** China Telecom Wanwei Information Technology Co., Ltd.

**Country**: China

**Contact Person(s)**:

BinBin Wen, 173735562@qq.com,

ShaoJie Zhang, 18919821844@189.cn

## 1 Use Case Summary Table

| Item | Details |
|---|---|
| Category | Smart Agriculture |
| Problem Addressed | Over reliance on individual experience, shortage of technical personnel, difficulties in promoting advanced agricultural technologies, low agricultural informatization level, limited service coverage, inconvenient access to technical information. |
| Key Aspects of Solution | Artificial Intelligence (AI) - powered agricultural Question & Answer assistant that leverages Interactive Natural Language Processing (NLP) for real-time planting guidance and knowledge dissemination. It also includes a pest and disease identification module using image-based AI recognition with over 86% accuracy, offering tailored treatment plans. Additionally, the platform integrates price forecasting capabilities through NLP and time-series analysis to deliver reliable market trend predictions. |
| Technology Keywords | AI/ML, LLM, Computer Vision, Natural Language Processing (NLP), Data Analytics. |
| Data Availability | Private Data (e.g., proprietary agricultural production records, pest/disease image databases from partner farms, historical market price datasets). Sources not publicly disclosed. |
| Metadata (Type of Data) | Text (for Q&A, price analysis input/output), Visual (for pest/disease identification images), Structured Data (for price analysis and prediction). |
| Model Training and Fine-Tuning | CV Models: Pre-trained YOLO fine-tuned on pest/disease datasets. Large Language Model (LLM): Pre-trained Qwen2.5-32B for Q&A and price trend analysis. Time-Series Forecasting: LSTM models for market predictions. |
| Testbeds or Pilot Deployments | AI - powered agricultural assistant offering technical guidance, price forecasting, and decision support to farmers through natural language interaction [4] |

# 2 Use Case Description

## 2.1 Description

Divine Farmer Agricultural AI Model is trained on massive agricultural production data to address current challenges in agricultural practices, including over-reliance on individual experience, shortage of technical personnel, and difficulties in promoting advanced agricultural technologies. It provides practitioners with intelligent agricultural services such as agricultural Q&A, pest and disease identification and prevention, and agricultural product price analysis and forecasting.

Agricultural Q&A: Acts as an intelligent agricultural assistant, offering planting guidance, farming recommendations, and technical knowledge through interactive Q&A. Eliminates geographical and temporal constraints, significantly improving the coverage and efficiency of agricultural technical services. Effectively resolves issues such as low agricultural informatization levels, limited service coverage, and inconvenient access to information.

Pest and Disease Identification and Prevention: Rapidly and accurately identifies pest and disease types from user-uploaded images (accuracy rate exceeding 86%). Generates targeted prevention and treatment plans, providing precise and scientific management tools for agricultural production. Protects crop health and ensures sustainable growth.

Agricultural Product Price Analysis and Forecasting: Delivers price analysis and market trend predictions through natural language Q&A. Helps users better grasp market dynamics, optimize resource allocation, and enhance the efficiency and profitability of the agricultural supply chain.

### Related information about our model:

Agricultural technology assistant: Q&A accuracy rate of over 90%. Disease and pest identification and prevention: Supports the identification of over 20 types of crops, 70 types of pests, and 136 types of diseases, with an accuracy rate of over 85%. Analysis and Trend Prediction of Agricultural Product Prices: The predicted data covers 441 types of agricultural products from 34 categories across 31 provinces, 141 cities, and 219 farmer's markets, with an accuracy rate of over 46% in predicting price trends.

## 2.2 Benefits of the use case

Alleviate rural poverty by providing farmers with precise cultivation guidance and market price forecasts. This helps optimize farming practices, increase yields, and boost income through informed decisions on sowing, irrigation, and harvesting.

Support food security through early pest and disease detection, minimizing crop losses and ensuring stable food supplies. The model enables timely interventions with advanced monitoring and alert systems.

Improve health outcomes by reducing reliance on chemical pesticides, thereby lowering harmful residues in food and limiting health risks for both farmers and consumers.

Bridge the technological gap in underdeveloped regions by offering equal access to smart farming tools and agricultural knowledge, helping reduce rural-urban disparities and promoting inclusive development.

Encourage sustainable resource use by intelligently managing inputs like water, fertilizers, and energy, enhancing efficiency and minimizing environmental waste.

Contribute to climate change mitigation by supporting precision agriculture techniques that reduce greenhouse gas emissions, such as optimized fertilization planning based on real-time environmental data.

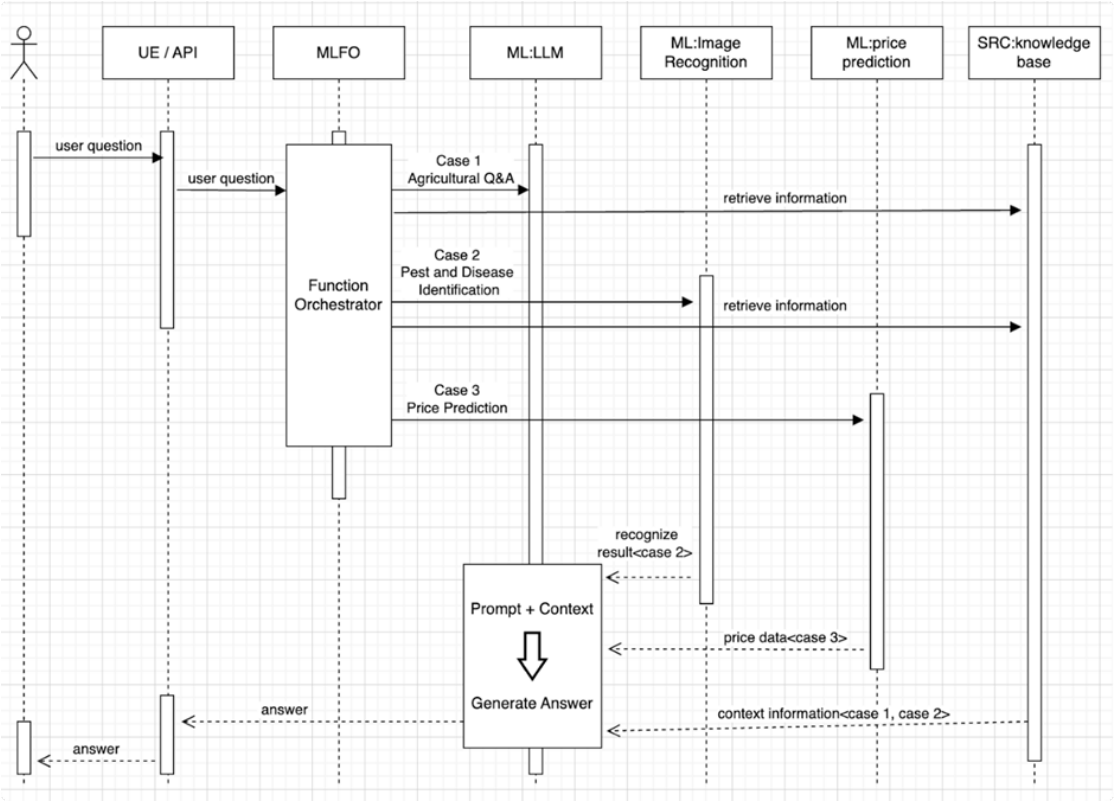
### 2.3 Future Work

- Data collection (Agricultural knowledge base, crop and plant disease/pest images, market information data, etc).
- Model development (Enhance the accuracy of Q&A and pest/disease identification; Add agricultural policy Q&A and interpretation, soil testing and formulated fertilization, etc.)

## 3 Use Case Requirements

- **REQ-01:** It is required to provide an interactive question-and-answer (Q&A) interface for addressing agricultural topics.
- **REQ-02:** It is required to deliver planting guidance through the Q&A interface to assist farmers in crop planning.
- **REQ-03:** It is required to offer farming recommendations via the Q&A interface for improved agricultural practices.
- **REQ-04:** It is required to make agricultural technical knowledge accessible through the Q&A interface to support informed decision-making.
- **REQ-05:** It is required to enable users to upload images for pest and disease identification.
- **REQ-06:** It is required to identify pest types accurately from user-uploaded images.
- **REQ-07:** It is required to identify disease types accurately from user-uploaded images.
- **REQ-08:** It is required to achieve an identification accuracy rate exceeding 86% for pests and diseases detected from images.
- **REQ-09:** It is required to generate treatment plans for pests and diseases identified through image analysis.
- **REQ-10:** It is recommended to provide agricultural product price analysis through a natural language Q&A interface.
- **REQ-11:** It is recommended to offer market trend predictions for agricultural products via the Q&A interface.

## 4 Sequence Diagram




## 5 References

[1] Li, Xiaoyan, Ying Zhang, and Hui Wang. 2022. "A Knowledge Graph-Based Agricultural Q&A System for Intelligent Farming Assistance." IEEE Transactions on Agri-Food Electronics. https://doi.org/10.1109/TAFE.2022.1234567.

[2] Chen, Li, et al. 2023. "Deep Learning for Crop Pest and Disease Identification: A Multi-Modal Image Analysis Framework." Computers and Electronics in Agriculture 198: 107023. https://doi.org/10.1016/j.compag.2023.107023.

[3] Zhang, Qi, and Rui Liu. 2020. "Machine Learning for Agricultural Price Forecasting: A Time-Series and NLP Hybrid Model." ACM Transactions on Intelligent Systems and Technology 11, no. 4: 1–22. https://doi.org/10.1145/3410352.

[4] SNYH. n.d. Accessed June 19, 2025. https://snyh.ilaoxiang.com.cn.

# Use Case 5: Edge AI and IoT-Powered System for Early Tomato Disease Detection and Smart Control

**Organization:** Mbeya University of Science and Technology (MUST)

**Country:** Tanzania

**Contact Person(s)**:

Yona H. Mgongolwa, yonamgongolwa@gmail.com,

Mr. Zacharia Mzurikwao, zacerbonnie@gmail.com

## 1 Use Case Summary Table

| Category | Smart Agriculture |
|---|---|
| The Problem Addressed to be addressed | Delayed detection of tomato plant diseases leads to significant yield reduction and increased economic losses for farmers. |
| Key aspects of the solution | Deployment of Edge-Artificial Intelligence (AI) for real-time disease detection, integrated with Internet of Things (IoT) - based smart monitoring to enable timely and informed crop management decisions. |
| Technology keywords | Edge-AI, IoT, Tomato Leaf Diseases, YOLOv9, Smart Agriculture |
| Data Availability | Open-source image and sensor datasets sourced from platforms such as Kaggle and Plant Village. |
| Metadata (type of data) | Visual data (images of tomato leaves), Sensor data (temperature, humidity) collected in real farming environments. |
| Model Training and fine-tuning | YOLOv9 deep learning model trained for accurate tomato leaf disease classification (achieving 98.6% accuracy); optimized for low-power edge and mobile deployment. |
| Testbeds or pilot deployments | Pilot tested on the demonstration farm at Mbeya University of Science and Technology (MUST). |

## 2 Use Case Description

### 2.1 Description

Agriculture remains a cornerstone of Tanzania's economy, particularly in the Southern Highlands region, which includes Mbeya, Njombe, Rukwa, Ruvuma, and Songwe. Among the major crops cultivated, tomatoes play a critical role in both economic stability and food security. However, tomato production in these areas faces a persistent threat from diseases such as early blight,

late blight, and septorial leaf spot, which cause substantial yield losses and directly affect the livelihoods of smallholder farmers.

In 2021, tomatoes accounted for more than 30% of Tanzania's total agricultural output, yet these diseases resulted in crop losses of up to 40% annually in the Southern Highlands. Smallholder farmers, who make up a significant portion of the agricultural workforce, are especially vulnerable. Traditional manual inspection methods are slow, imprecise, and inadequate for early intervention, underscoring the urgent need for a real-time, AI-powered disease detection system. This solution aims to focus on the development of an Edge AI and IoT-powered system for early tomato disease detection and smart control. The system integrates advanced technologies to deliver accurate, real-time identification and notification of tomato leaf diseases.

### Specific Objectives:

- To design and simulate a deep learning model capable of detecting and distinguishing between healthy and diseased tomato leaves.
- To develop a real-time alert system that notifies farmers when disease symptoms are detected.
- To build a mobile application that delivers insights and guidance based on model outputs.

At the core of this solution is the YOLOv9 deep learning model, optimized for deployment on edge devices such as Raspberry Pi. This architecture allows for real-time image analysis directly on the farm, removing the dependency on continuous internet access an essential feature for rural agricultural environments.

By enabling early and accurate disease detection, this AI-IoT solution supports smallholder farmers in increasing crop yields, reducing losses, and minimizing unnecessary pesticide use. The real time alerts and mobile based insights ensure that even farmers in remote areas can take timely action, ultimately enhancing food security, promoting sustainable agriculture, and improving farmer livelihoods through precision farming.

## 2.2 Benefits of the use case

Improve smallholder farmers' livelihoods by reducing crop losses caused by tomato diseases. Through early detection and targeted disease management, the solution minimizes income loss and supports more stable, resilient agricultural practices.

Strengthen food security by increasing tomato crop yields and ensuring a more reliable food supply. By enabling timely and accurate disease intervention, the system contributes to sustained agricultural productivity and helps combat hunger in rural communities.

Support ecosystem conservation and biodiversity protection by promoting sustainable disease control techniques. The reduction in unnecessary pesticide use helps maintain soil health and reduces harm to beneficial organisms, aligning with broader goals of sustainable land use and environmental stewardship.

## 2.3 Future Work

Future work includes expanding the dataset to cover a wider range of tomato varieties and environmental conditions to improve model robustness. Integrating drone based monitoring

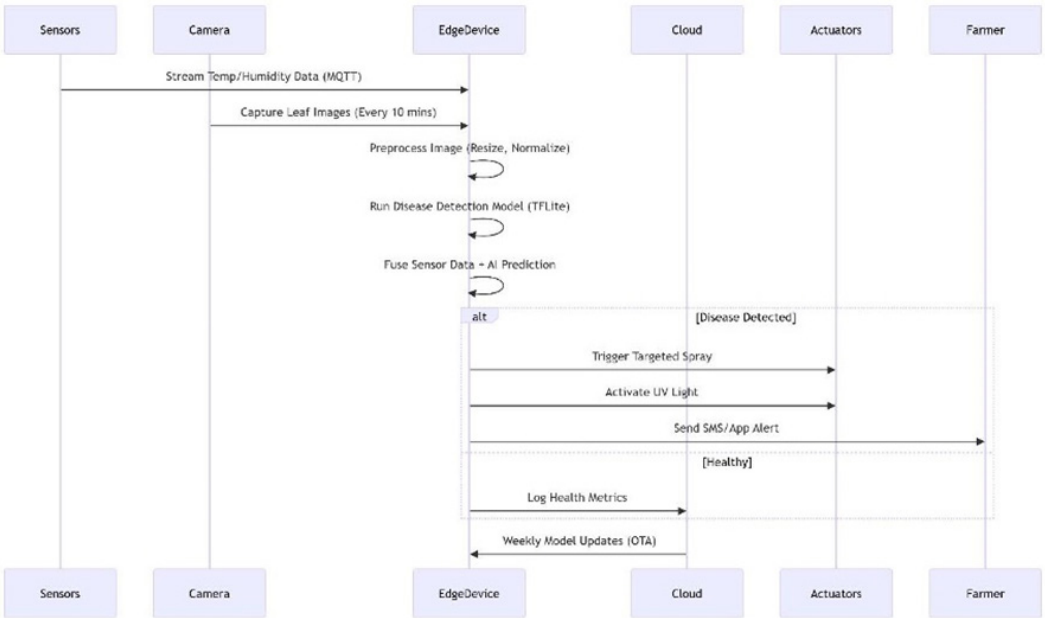
is planned to enhance large scale field surveillance and complement sensor based ground monitoring systems.

Further efforts will focus on optimizing AI models for faster performance on edge devices to support real-time decision making. Potential areas for scale up include forming partnerships for broader deployment and extending the system to support multi crop disease detection for wider agricultural impact.

## 3 Use Case Requirements

- **REQ-01**: It is required to collect high resolution images of tomato leaves for accurate disease detection and model training.
- **REQ-02**: It is required to deploy edge computing hardware such as Raspberry Pi to enable real time image processing on-site.
- **REQ-03**: It is required to install IoT sensors to monitor environmental conditions, specifically temperature and humidity, in tomato fields.
- **REQ-04**: It is required to develop a mobile application equipped with a camera interface and alert system to provide real-time notifications to farmers.
- **REQ-05**: It is required to ensure stable power supply and uninterrupted internet connectivity to support continuous data transmission and system updates.

## 4 Sequence Diagram




## 5 References

[1] Kaggle. n.d. Accessed June 19, 2025. https://www.kaggle.com/.

[2] Singh, Vikram, Nitin Sharma, and Sandeep Singh. 2020. "A Review of Imaging Techniques for Plant Disease Detection." *Artificial Intelligence in Agriculture* 4: 229–42. https://doi.org/10.1016/j.aiia.2020.10.002.

[3] Subeesh, Anil, and Chaitanya R. Mehta. 2021. "Automation and Digitization of Agriculture Using Artificial Intelligence and Internet of Things." *Artificial Intelligence in Agriculture* 5: 278–91. https://doi.org/10.1016/j.aiia.2021.11.004.

[4] Thangaraj, R., S. Anandamurugan, P. Pandiyan, and V. K. Kaliappan. 2022. "Artificial Intelligence in Tomato Leaf Disease Detection: A Comprehensive Review and Discussion." *Journal of Plant Diseases and Protection* 129, no. 3: 469–88. https://doi.org/10.1007/s41348-021-00500-8.

[5] Nobel, S. M. N., M. Afroj, M. M. Kabir, and M. F. Mridha. 2024. "Development of a Cutting-Edge Ensemble Pipeline for Rapid and Accurate Diagnosis of Plant Leaf Diseases." *Artificial Intelligence in Agriculture* 14: 56–72. https://doi.org/10.1016/j.aiia.2024.10.005.

# Use Case 6: Intelligent Soil Analyzer and Crop Fertilizers Advisory System

**Organization**: Mbeya University of Science and Technology (MUST)

**Country**: Tanzania

**Contact Person**:

Projestus Protas projestusprotas601@gmail.com

## 1 Use case summary Table

| Category | Smart Agriculture |
|---|---|
| The problem to be addressed | Farmers choose crops to plant and fertilizer to use in the farm basing on their own understanding and historical practices of previous farmers. This leads to inefficient resource use and low crop yields. |
| Key aspects of the solution | This system will provide accurate recommendations and suggestions to farmers about a suitable crop to cultivate in the farm and amount of fertilizer needed basing on nutrients level in the farm. |
| Technology keywords | Intelligent soil analyzer, Crop fertilizer advisory system, IoT sensors, Machine learning, Remote server and Sustainable agriculture. |
| Data availability | Private (Collected on site by IoT sensors) and Public ( Collected from Kaggle archive platform for training datasets) |
| Metadata (type of data) | Numerical data (NPK values, soil pH , soil moisture, rainfall and temperature) and Text data (Crop type, Fertilizer type) |
| Model Training and fine-tuning | Random forest classifier is going to be the main machine learning algorithm to processes and analyze the collected data |
| Testbeds or pilot deployment | Plantix. n.d [3]<br>CropIn. n.d [4] |

## 2 Use case description

### 2.1 Description

Agriculture is backbone of the economy, but farmers face issues like poor crop selection, soil degradation, inefficient fertilizer use, and nutrient deficiencies. Current farming methods don't effectively address these challenges. There is a need of system that provides real-time insights

into soil and nutrient contents, helping farmers make informed decisions and improve their productivity.

Existing system suggests only suitable crop to plant, as a result, farmers may follow the recommended crop choices but experience low yields due to a lack of essential nutrients in the soil. Furthermore, these systems do not provide fertilizer guidelines and crop planning scheme leading to nutrient imbalances and potential environmental harm.

The objectives and aims of this use case are:

- To design an IoT-based soil analysis system that can measure key soil parameters.
- To develop Machine-learning algorithms (Random Forest Classifier) capable of processing and analyzing soil data.
- To develop a fertilizer system that identifies the type of crop being cultivated, analyzes soil nutrients, and suggests suitable fertilizers.
- To create a user-friendly dashboard interface that allows farmers to easily access soil analysis reports and recommendations.

This innovative approach operates by collecting soil data through sensor **JXCT Soil NPK Sensor RS485** that measure key soil parameters, including pH, moisture, temperature, and nutrient levels (N, P, K). The collected data are preprocessed by a microcontroller and then transmitted to a remote server for analysis via Wi-Fi. The remote server hosts Database and machine-learning algorithms that analyze soil quality and determine the nutrient requirements of a given crop or suggest the most suitable crop to cultivate. Based on this analysis, the system generates actionable recommendations for farmers, including the type and quantity of fertilizers needed, suitable crop suggestions, and crop rotation schedules.

**Partners**

- Tanzania Agricultural Research Institute (TARI).
- Tanzania Fertilizer Regulatory Authority (TFRA) Sokoine University of Agriculture (SUA)

## 2.2 Benefits of the solution

1. To improve and boosts production and farmer income. This help farmer to improve yield by providing advanced fertilizer advice based on real time soil data. This will enable better productivity even in smallholder farmer and farmer income.
2. To promote optimized use of agricultural resources.
3. To reduce fertilizer overuse. Optimizing fertilizer use reduces greenhouse gas emissions also to support agriculture by improving resilience to weather patterns.
4. Preserves soil health and biodiversity. Promotes soil conservation and biodiversity by reducing harmful agricultural practices.

## 2.3 Future work

Currently, soil analyzer systems only suggest suitable crops based on basic soil parameters such as pH and moisture. However, these systems fail to account for critical factors such as nutrient deficiencies and fertilizer requirements. As a result, farmers may follow the recommended crop choices but experience low yields due to a lack of essential nutrients in the soil. We

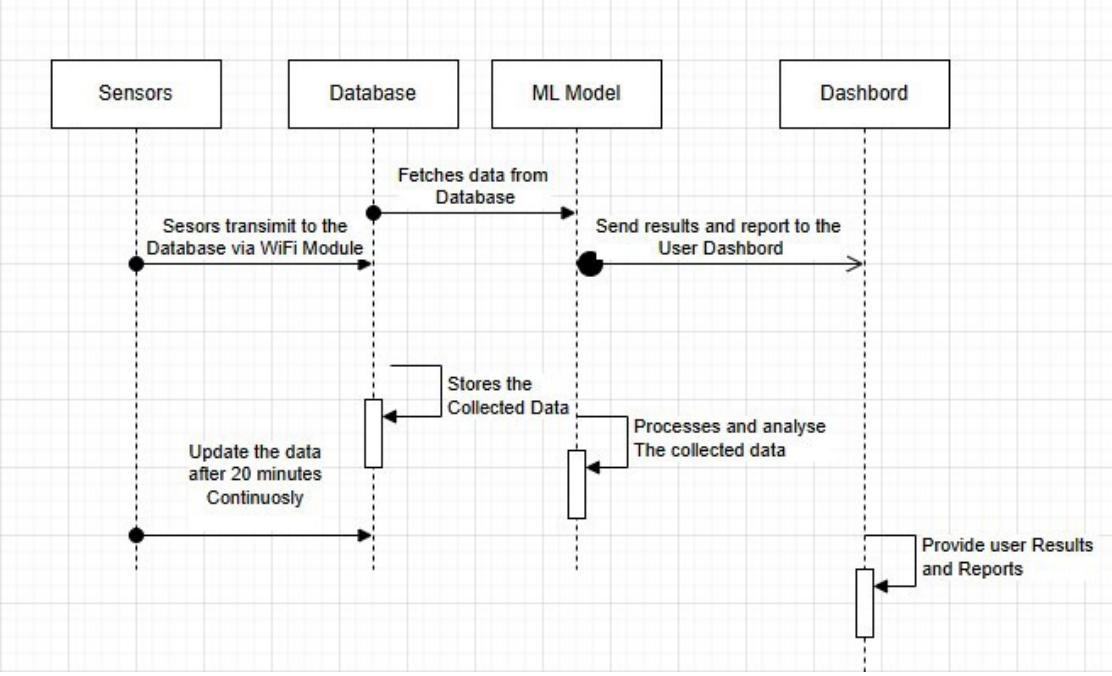
address these challenges by providing fertilizer guidelines and crop planning scheme leading to nutrient balances and reduction of potential environmental harm in the farm.

We propose this system(AI based) to leverage soil analysis for precision fertilization and crop rotation strategies that enhance crop production, conserving resources and improving soil sustainability for the future use.

## 3 Use case requirements

- **REQ-O1**: Soil data collection- Soil sensors must capture pH, moisture, temperature and NPK values continuously from the farm field.
- **REQ-O2**: Transmission- Collected data should be transmitted securely via Wi-Fi to the remote sever (Database )
- **REQ-O3**: Machine learning model-The model should fetch the collected data from database for processing and analysing to provide recommendations and advises.
- **REQ-O4**: User friendly interface- To interact with system to see reports and results

## 4 Sequence diagram




## 5 References

[1] Abinaya, P., A. Faaiza, S. Tharshni, G. Vishali, M. Aravindan, and D. G. Santhi. "Survey on Soil Micro-Nutrients Analysis and Crop Recommendation System in Smart Agriculture." *International Research Journal on Advanced Engineering Hub (IRJAEH)* 2, no. 03 (March 2024): Article no. 03. https://doi.org/10.47392/IRJAEH.2024.0066.

[2] Sankareswari, K., and G. Sujatha. "Crop and Fertiliser Recommendation System for Sustainable Agricultural Development." *In Intelligent Robots and Drones for Precision Agriculture*, edited by S. Balasubramanian, G. Natarajan, and P. R. Chelliah, 327–49. Cham: Springer Nature Switzerland, 2024. https://doi.org/10.1007/978-3-031-51195-0_16.

[3] Plantix. n.d. Accessed June 19, 2025. https://www.plantix.net/.

[4] CropIn. n.d. Accessed June 19, 2025. https://www.cropin.com/.

International Telecommunication Union
Place des Nations
CH-1211 Geneva 20
Switzerland

ISBN 978-92-61-40961-6

Published in Switzerland
Geneva, 2025

Photo credits: Adobe Stock