Alliance for AI, IoT and Edge
Continuum Innovation

AI, IoT and Edge Continuum impact and relation on 5G/6G: enabling technologies and challenges

Release 5.0

AIOTI WG Standardisation

20 October 2025

Executive Summary

This report highlights several IoT and Edge Computing vertical domain use cases collected by the Alliance for AI, IoT and Edge Continuum Innovation (AIOTI) and determines the specific requirements they impose on the underlying 5G/6G network infrastructure. These use cases and requirements can be used by Standards Developing Organizations (SDOs), such as 3GPP, ITU-T, ISO, and IEEE as requirements for automation in vertical domains focusing on critical communications.

In addition to these use cases also emerging topics in the area of (Beyond) 5G technology are as well introduced.

The Release 2.0 of this report included 6 additional use cases in the areas of: (1) use of drones, (2) 5G cloud-RAN, (3) Health-Critical Remote Operations, (4) preliminary 6G use cases.

The Release 3.0 of this report included 6 additional use cases in the area of Edge-Cloud Orchestration in the Section 2.13.

The Release 4.0 of this report included 14 additional use cases in the areas of: (1) Digital Twin, (2) autonomous urban transportation, (3) critical Infrastructure support applications (smart health and connected vehicles), (4) preliminary 6G use cases, (5) use of drones, (6) smart manufacturing and automation, (7) service trust and liability management, (8) Edge-Cloud orchestration and (9) smart agriculture.

The Release 5.0 of this report, compared to Release 4.0, introduced the following changes:

1. the use cases, developed in projects that are: (1) completed before 2025 and (2) were included in Release 4.0 of the report, are not anymore included in this release (Release 5.0) of the report, but they can be retrieved via Section 2 of AI, IoT and Edge Continuum impact and relation on 5G/6G: enabling technologies and challenges Release 4.0.

   - in Section 3, added a subsection on Key AI concepts

   - in Section 2, a list of additional 12 use cases are introduced in the areas of (1) Smart Manufacturing and Automation, (2) Preliminary 6G use cases and (3) Edge-Cloud orchestration.

Table of Content

Table of Figures

List of Tables

Abbreviations

2D	Two Dimensional
3GPP	3rd Generation Partnership Project
4G	4th Generation
5G	5th Generation
5G-NPN	5G Non-Public Network
ABS	Anti-lock Braking System
ACL	Access Control Lists
ADApp	Autonomous Driving Application
AF	Application Function
AGV	Automated Guided Vehicle
AI	Artificial Intelligence
AIOTI	Alliance for IoT Innovation
App	Application
AR	Augmented Reality
AS	Application Server
ASF	Authentication Server Function
AVP	Automated Valet Parking
BDA	Big Data Analytics
BICMOS	Bipolar Complementary Metal—Oxide-Semiconductor
BLE	Bluetooth Low Energy
BMS	Building Management System
BVLOS	Beyond Vision Line of Sight
C-ITS	Cooperative-Intelligent Transportation System
CAD	Connected and Automated Driving
CAGR	Compound Annual Growth Rate
CAM	Cooperative Awareness Message
CAPEX	Capital Expenditure
CC	Cloud Computing
CCAM	Connected and Automate Mobility
CNN	Convolutional Neural Network
CPS	Cyber–Physical Systems
CRAN	Cloud Radio Access Network
CSS	Car Sharing Service
D2X	Device to everything
DoF	Degree of Freedom
DoS	Denial-of-Service
DTw	Digital Twins
EEW	Early Earthquake Warning
eMBB	Enhanced Mobile Broadband
EPON	Ethernet Passive Optical Network
ESP32	Espessif Systems Processor 32
ETSI	European Telecommunication Standardisation Institute
FFT	Fast Fourier Transform

6

FL	Federated Learning
FoF	Factories of the Future
GDPR	General Data Protection Regulation
GPS	Global Positioning System
GSM	Global System for Mobile communications
HPC	High Performance Computing
I&M	Inspection & Maintenance
IDS	Intrusion Detection System
IEM	Interaction of Employees and Machines
IIoT	Industrial Internet of Things
IoRT	Internet of Robotic Things
IoT	Internet of Things
IP	Internet Protocol
ITS	Intelligent Transportation System
LDM	Local Dynamic Map
LoRa WAN	LoRa Wide Area Network
LOS	Line Of Sight
LP-WAN	Low Power Wide Area Network
LTE	Long Term Evolution
LTE-V2X	LTE Vehicle to Everything
MCU	MicroController Unit
MEC	Multi-access Edge Computing
MES	Manufacturing Execution System
MIMO	Multiple Input, Multiple Output
ML	Machine Learning
mMTC	Machine-Type Communications
MQTT	Message Queuing Telemetry Transport
MUD	Manufacturer Usage Description
NACF	Network Access Control Function
NB-IoT	Narrowband IoT
NFR	Network Function Registry
NFV	Network Function Virtualisation
NoLOS	Non-Line of Sight
NoSQL	Not only Structured Query Language
NPN	Non-Public Network
NR	New Radio
NSSF	Network Slice Selection Function
NTN	Non-Terrestrial Networks
OBU	On-Board Unit
OEM	Original Equipment Manufacturer
OGC	Open Geospatial Consortium
OPEX	Operational Expenditure
OT	Operation Technology
OTA	Over The Air

PCF	Policy Control Function
RP-an	reference point between AN and NACF
RP-au	reference point between AN and UPF
RP-tn	reference point between UE and NACF
RP-ud	reference point between UPF and data network
RSU	Road Side Unit
RUL	Residual Useful Life
SAS	Service Alerting System
SCADA	Supervisory Control and Data Acquisition
SDO	Standards Developing Organization
SDO	Standards Developing Organizations
SME	Small Medium Enterprise
SMF	Session Management Function
SNS JU	Smart Network and Services Joint Undertaking
SOI	Silicon-On-Insulator
TC	Technical Committee
TCP	Transmission Control Protocol
TIoT	Tactile Internet of Things
TMC	Traffic Management Center
TSC	Time Sensitive Communication
TSN	Time-Sensitive-Networking
UAS	Unmanned Aerial System
UAV	Unmanned Aerial Vehicle
UE	User Equipment
UMTS	Universal Mobile Telecommunication System
UPF	User Plane Function
uRLLC	Ultra-reliable and Low-latency Communications
USM	Unified Subscription Management
UTM	Unmanned Traffic Management system
V2V	Vehicle to Vehicle
vApp	Vertical Application
VR	Virtual Reality
VRU	Vulnerable Road Users
WAVE	Wireless Access in Vehicular Environments
WiMAX	Worldwide Interoperability for Microwave Access
WLAN	Wireless Local Area Network
WSDN	Wireless Software Defined Network
XML	Extensible Markup Language

1. Introduction

As emphasized in reports published by AIOTI, the IoT is projected to consist of 50 billion devices by 2020 [Evans11] ranging from connected temperature sensors to autonomous vehicles. The vast scope of different device types from different verticals corresponds with highly diverse requirements for the communication infrastructure. While battery-driven sensors need a highly energy efficient communication technology, industrial IoT applications call for ultra-reliable connections with a minimum latency.

Important to mention that the ubiquitous nature of IoT devices has triggered a change to the models of managing and controlling the flow and transmission of data. The new concepts are moving from the widespread use of cloud-based infrastructure models, which are dominated by leading Internet companies, towards IoT edge mesh distributed processing, low latency, fault tolerance and increased scalability, security, and privacy.

As of today, these diverse requirements are covered by several wireless communication technologies (e.g. WLAN, Sigfox®, ZigBee, LoRaWAN, NB-IoT) which all have their specific strengths and weaknesses and that are making the IoT somewhat of a "rag rug".

This is where the 5G and beyond 5G becomes to be relevant, with its highly flexible architecture designed to be adaptable to almost any use case in the IoT space using advanced techniques like network slicing and NFV, see e.g., [Networld2020-SRIA[1]], [5GPPP-Vision], [5GPPP-verticals]. By offering a unified communications platform for the IoT, 5G has the potential of being a catalyst for IoT growth – and vice versa.

The "IoT Relation and Impact on 5G" AIOTI report [AIOTI-IoT-relation-5G] focused on highlighting emerging topics and specific IoT vertical domain use cases and determine the specific requirements they impose on the 5G network infrastructure.

This report focuses on highlighting new emerging topics and specific IoT vertical domain use cases and determine the specific requirements they impose on 5G and as well beyond 5G network infrastructure. These use cases and requirements can be used by SDOs, such as 3GPP, ISO, ITU-T and IEEE as requirements for automation in vertical domains focusing on critical communications.

The Release 2.0 of this report includes 6 additional use cases in the areas of: (1) use of drones, (2) 5G cloud-RAN, (3) Health-Critical Remote Operations, (5) preliminary 6G use cases. In particular, the added use case cases are:

- Multi-tenant real time AI video/audio analytics
- AURORAL HEALTH PILOT for Strengthening Preparedness In Health-Critical Remote Operations
- Virtualized base station for 5G cloud-RAN
- Hexa-X 6G based Use cases
- Connectivity during crowded events use case, when drones are used
- An innovative fire detection pilot solution using 5G, Artificial Intelligence and drone technology.

---

[1] Networld2020 ETP has been renamed to NetwolrdEurope ETP, see: https://www.networldeurope.eu

The [Release 3.0](#) of this report also includes 6 additional use cases in the area of Edge-Cloud Orchestration, which are:

- CODECO P1: Smart Monitoring of the Public Infrastructure,
- CODECO P2: Vehicular Digital Twin for Safe Urban Mobility,
- CODECO P3: MDS across Decentralised Edge-Cloud,
- CODECO P4: Demand-side Management in Decentralized Grids,
- CODECO P5: Wireless AGV Control in Flexible Factories,
- CODECO P6: Automated Crownstone Application Deployment for Smart Buildings.

The [Release 4.0](#) of this report also includes 14 additional use cases in different areas, which are:

- ERATOSTHENES: Connected Vehicles,
- ERATOSTHENES: Smart Health,
- EVOLVED-5G: "Efficiency in FoF Operations with Novel Predictive Maintenance applied on Digital Factory Twin,
- 5G-VICTORI: UC #1.1: Enhanced Mobile Broadband under High Speed Mobility,
- 5G-VICTORI: UC #2: Factories of the Future,
- 5GMETA: Driving Safety & Awareness,
- 5G COMPLETE: Example: UC#4: Advanced Surveillance/Physical Security Service,
- 5G COMPLETE: UC#3: 5G Wireless Transport services with MEC capability provided to NOs
- RISE-6G: Control for RIS-based localisation and sensing,
- AI@EDGE: Edge AI assisted monitoring of linear infrastructures using drones in BVLOS operation,
- 5G-INDUCE: Drone assisted network performance and coverage monitoring for industrial infrastructures,
- 5G-INDUCE: ML-Supported Edge Analytics for Predictive Maintenance,
- COMMECT: Monitoring of Pest Insect Traps,
- COMMECT: Securing crops and equipment.

The Release 5.0 of this report, compared to Release 4.0, introduced the following changes:

2. the use cases, developed in projects that are: (1) completed before 2025 and (2) were included in Release 4.0 of the report, are not anymore included in this release (Release 5.0) of the report, but they can be retrieved via Section 2 of [AI, IoT and Edge Continuum impact and relation on 5G/6G: enabling technologies and challenges Release 4.0](#).
- in Section 3, added a subsection on Key AI concepts
- in Section 2, a list of additional 12 use cases are introduced, see:
    - Factory-in-a-Box: Multiprotocol Beyond-5G Testbed for Smart Manufacturing
    - 6G-BRICKS (6G- Building Reusable testbed Infrastructures for validating Cloud-to-device breakthrough technologies) Use Cases
        - Use Case 1: Metaverse as an enabler of a Modern Workplace
        - Use Case 2: 6G applications for Industry 4.0
            - PoC 1: Autonomous robots in Industry 4.0
            - PoC 2: AR inspection of Industry 4.0 digital twin on site

- 6G-DALI (6G DAta and ML operations automation via an end-to-end AI framework) Use Cases
  - Use Case 1: Data Management and Experiment-on-Demand
  - Use Case 2: AIaaS for CDN Apps via Cross-Testbed Decentralized MLOps
  - Use Case 3: DTT and RLOps for Large and Medium-Scale Experiments
- 6G-INTENSE (Intent-driven NaTive AI architecturE supporting Compute-Network abstraction and Sensing at the Deep Edge) Use Cases
  - Use Case 1: Distributed Continuum towards Pervasive Computing
    - Use Case 1 - Scenario 1: Pervasive Computing in a distributed continuum
    - Use Case 1 - Scenario 2: Edge Intelligence and Compute Interconnection
  - Use Case 2: Metaverse
    - Use Case 2 - Scenario 1: Joint Communication and Sensing for Optimal User Tracking in the Metaverse
    - Use Case 2 - Scenario 2: Fully Autonomous Metaverse FCAPS, sensing and Continuum Abstraction
- 2.11.6 SUNRISE-6G (SUstainable federatioN of Research Infrastructures for Scaling-up Experimentation in 6G) Use Cases
  - Use Case 1 : Federated Metaverse
    - Use Case 1 - Scenario 1: Federation via Edge-Cloud capabilities and APIs
    - Use Case 1 - Scenario 2: Exposure and Consumption of Network Capabilities
  - Use Case 2: Collaborative Robotic Exploration of Mining Environments
  - Use Case 3: Mobile Industrial Assets aided by NTN Edge Cloud & Communications
    - Use Case 3 - Scenario 1: Multi-RAT over 5G NR and legacy NTN
    - Use Case 3 - Scenario 2: Satellite Edges for Decentralized Data Storages
    - Use Case 3 - Scenario 3: Migration of application context
- COGNIT - CCAM services enhanced with AI assisted CEI framework

2. Human Centric and Vertical Services and Use cases for Beyond 5G

This section describes the IoT vertical domain use cases that are being developed in IoT focused projects. Moreover, this section describes he specific requirements that these use cases impose on the underlying network infrastructure.

The use cases listed in this section have been described using the use case description template provided in Annex II.

2.1 Smart Manufacturing and Automation

5G supports communication with unprecedented reliability and very low latencies, and also massive IoT connectivity. This paves the way for numerous new use cases and applications in many different vertical domains, including the automotive, healthcare, agriculture, energy and manufacturing sectors. In manufacturing in particular, 5G may have a disruptive impact as related building blocks, such as wireless connectivity, edge computing or network slicing, find their way into future smart factories.

The fourth stage of the Industrial Revolution, also termed "Industry 4.0", is the next era in industrial production, aiming at significantly improving the flexibility, versatility, usability and efficiency of future smart factories.

Industry 4.0 integrates the Internet of Things (IoT) and related services in industrial manufacturing and delivers seamless vertical and horizontal integration down the entire value chain and across all layers of the automation pyramid [KaWa13] – here named Industrial IoT (IIoT). Connectivity is a key component of Industry 4.0 and will support the ongoing developments by providing powerful and pervasive connectivity between machines, people and objects. Moreover, wireless communication, and in particular 5G, is an important means of achieving the required flexibility of production, supporting new advanced mobile applications for workers, and allowing mobile robots and autonomous vehicles to collaborate on the shop floor – these being just a few examples.

Some of the target key performance indicators of 5G as specified by the International Telecommunications Union (ITU) are summarized in **Figure 1** (cf. [ITU-R M.2410-0]).

In order to support the three service types defined above and the diverse requirements of the anticipated 5G use cases by a common cellular infrastructure, network slicing, a new concept introduced in 5G, will allow simultaneous but isolated provisioning of diverse services by the same network infrastructure.
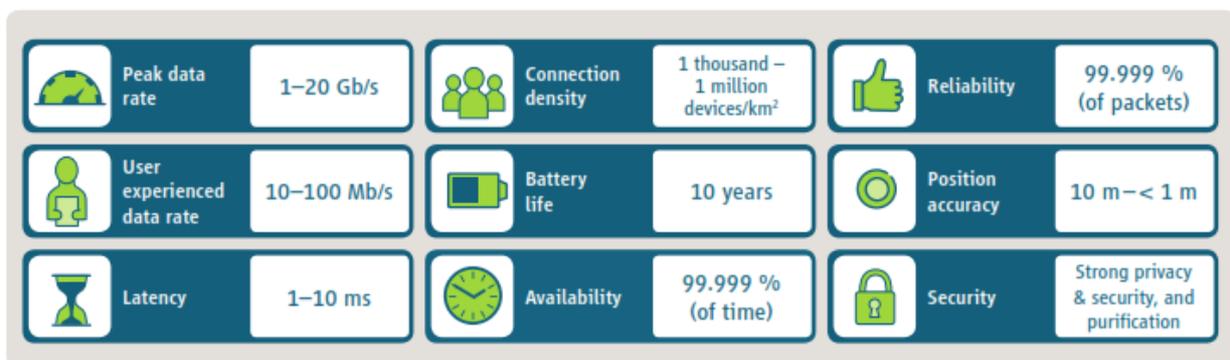
Figure 1: Selected target key performance indicators of 5G according to ITU-R (cf. [ITU-R M.2410-0])

Industry 4.0 and the Role of 5G

The fourth stage of the Industrial Revolution, also termed "Industry 4.0", is the next era in industrial production, aiming at significantly improving the flexibility, versatility, usability and efficiency of future smart factories. Industry 4.0 integrates the Internet of Things (IoT) and related services in industrial manufacturing and delivers seamless vertical and horizontal integration down the entire value chain and across all layers of the automation pyramid [KaWa13]. Connectivity is a key component of Industry 4.0 and will support the ongoing developments by providing powerful and pervasive connectivity between machines, people and objects. Moreover, wireless communication, and in particular 5G, is an important means of achieving the required flexibility of production, supporting new advanced mobile applications for workers, and allowing mobile robots and autonomous vehicles to collaborate on the shop floor – these being just a few examples.

5G Roadmap

The 3GPP began work on the specification of 5G in early 2017. The standardization work has been divided into two major phases: standardization of the fundamental 5G building-blocks has already been completed in June 2018 (Release 15), and further enhancements added by the end of 2019 (Release 16). According to 3GPP SA2 the Release-17 work made good progress, which most of the study items are over 95% complete. The study focus related to IIoT is on enhanced support of standalone non-public networks "SNPN" (TR23.700-07) and on enhanced support of Industrial Internet of Things related to Time Sensitive Communication (TSC) (TR23.700-20) including enhancements for support of deterministic applications etc. to IEEE Time-Sensitive-Networking (TSN) which is supported by 5G-ACIA work items for manufacturing industries.

Looking ahead to 2026, digitalization revenues from 5G for ICT players are estimated to exceed 1,200 billion USD, of which approximately 234 billion USD is accounted for by the corresponding vertical manufacturing [ErLi17]. In business terms, this constitutes an incredibly large and fast-growing market.

2.1.1 Factory-in-a-Box: Multiprotocol Beyond-5G Testbed for Smart Manufacturing

2.1.1.1 Description

As the name implies, additive manufacturing creates objects by adding material layer by layer to achieve precise geometries. The medical sector—with its need for customizable, biocompatible, and sterilizable plastic and metal components—is one of the most promising fields for 3D-printing applications. On average, printing an object takes from thirty minutes to twelve hours, depending on complexity; more intricate builds may take longer. Print failures can postpone delivery of parts and extend patient wait times. By systematically exploring the effect of individual process parameters on print failure, we can optimize material usage and total print time. Our objective was to deliver a portable, standards-aligned "factory-in-a-box" that we could deploy and reconfigure rapidly while preserving deterministic, secure, and interoperable manufacturing across heterogeneous assets. We integrated 3GPP-based private 5G for ultra-reliable low-latency control and high-bandwidth media alongside Wi-Fi 6, LoRaWAN, and BLE for complementary workloads. Instead of OPC/AMQP, we operated a proprietary data-gathering layer built on Redis that ingested device telemetry, normalized it into a governed schema, and distributed it through Redis Streams/pub-sub and lightweight REST/WebSocket APIs. We anchored computation at the edge to host a digital twin and real-time analytics that fused high-rate vibration and thermal signals and triggered interventions when quality risks emerged during multi-hour printing and inspection operations. This use case was motivated by rising demand volatility, shortened product lifecycles, and recurrent supply-chain shocks, which make fixed, cabling-intensive lines costly to reconfigure. We applied the pattern in a working testbed and in targeted trials that combined additive manufacturing with inline inspection; these deployments demonstrated earlier anomaly detection, reduced scrap, and materially faster commissioning compared with traditional setups.

The business case was strengthened by a multi-stakeholder ecosystem: manufacturers and operations teams realized first-pass yield improvements, quicker changeovers; maintenance functions benefited from condition-based interventions and reduced unplanned downtime;

IT/OT engineering obtained a tractable integration surface via the Redis-backed hub; compliance and safety stakeholders gained stronger evidence chains; and end customers experienced shorter lead times, higher quality, and greater scope for late-stage customization. At a high level, the operation proceeded as follows. We planned the work cell and pre-provisioned network slices appropriate to control, video, and sensing. We onboarded devices with authenticated identities and attached least-privilege policies. Telemetry then flowed over the most suitable bearer into a back-end system tailored to each device's needs, where it was time-synchronized, normalized, and exposed to edge analytics and the digital twin. When anomalous signatures arose, we invoked MES actions to adjust set-points and, where required, paused or re-routed jobs while providing operator guidance. We persisted logs, labeled outcomes, retrained models, and concluded each run with auditable reports on performance, safety, and compliance to drive continuous improvement and replication across sites. The assembly line comprised the following integrated work cells: a UR5 robotic arm servicing all stations, a resin-mixing station that maintained material homogeneity, a medical-grade **Figure 3** 3D resin printer, a post-print curing station, an optical inspection microscope, a 3D digital-imaging inspection station, and a packaging station. **Figure 2** shows the initial virtual layout of the cells.
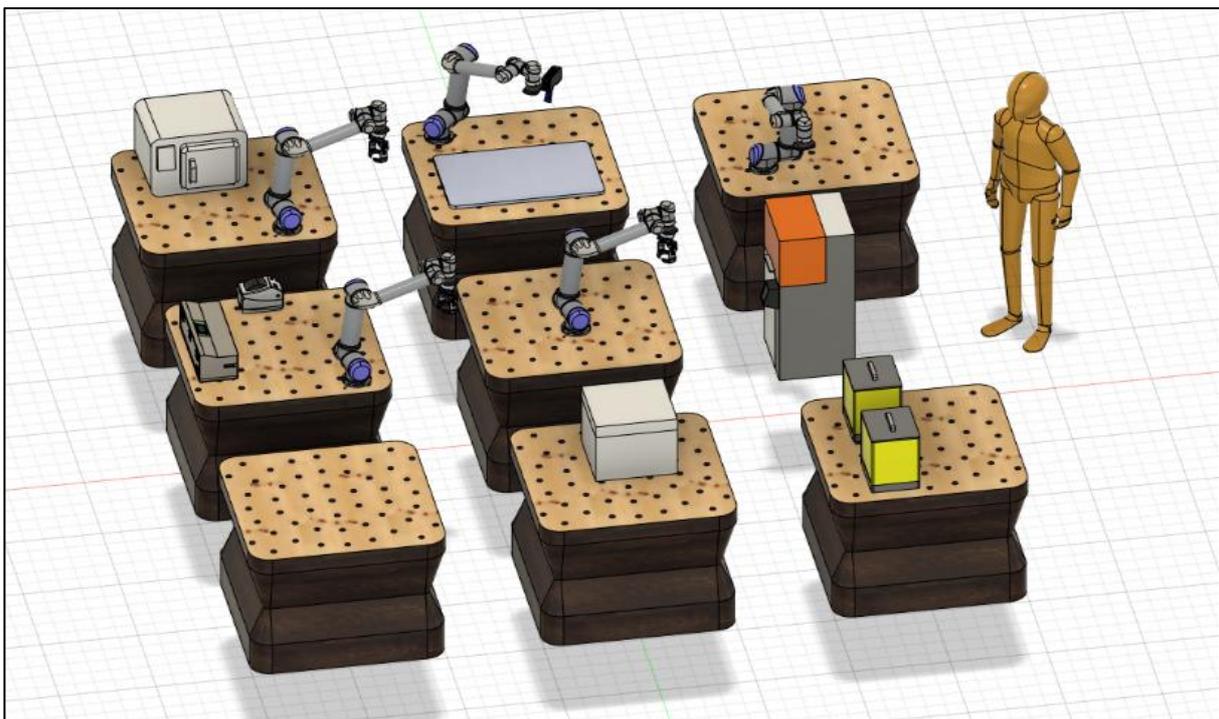
Figure 2 Initial virtual layout of the cells

2.1.1.2 Source

www.confirm.ie: CONFIRM Smart Manufacturing testbed "Factory-in-a-Box" lab setup (multiprotocol architecture, on-prem data centre, 3D printer predictive maintenance demo).

2.1.1.3 Roles and Actors

Roles:

- Clinical Sponsor / End User (medical provider). Specifies clinical requirements, provide patient consent and scan data, approve device designs, and receive finished parts for use or further validation.

- Vertical Industry Owner (medical-device manufacturer / plant owner). Owns production KPIs, GMP compliance, and release decisions; we fund deployment and oversee day-to-day operations.

- System Integrator & Testbed Operator. Architects the medical cell, integrates subsystems, operates the Redis-based data hub, and maintain the digital twin, edge analytics, and ML lifecycles.

- Communication Network Provider/Operator (private 5G). Assures non-public network services (RAN/5GC, slicing, on-prem UPF) to meet latency, reliability, and coverage targets.

- IoT/Automation Device Manufacturers. Supplies and supports the UR3 robotic arm, the **Figure 3** 3D printer, curing and cleaning equipment, sensors, and the EinScan 3D scanner; Also provides firmware, safety documentation, and maintenance.

- IoT/Edge/Platform Provider. Hosts governed storage, streaming, dashboards, model repositories, and APIs that expose telemetry from the Redis hub to enterprise systems.

- MES/ERP Provider. Orchestrates jobs, genealogy, and actuation (pause, re-route, set-point updates) and capture KPIs.

- Quality, Regulatory, and HSE. Offers validation and verification protocols, risk management, and audit trails; we ensure workplace and device safety.

- Cybersecurity & Audit. Enforces zero-trust policies, conduct assessments, review SBOMs, and preserve immutable logs for forensics.

- Logistics / Distribution. Packages, labels, and ships finished parts with full traceability to clinical sites.

Relationships between roles:

The industry owner contracts the communication network provider to deliver private-5G services under explicit SLAs and engages the MES/ERP provider to close the loop between analytics and production control. IoT/automation device manufacturers supply hardware and firmware under change-control, while the platform provider exposes governed data products from our Redis-based hub to enterprise applications. Quality, regulatory, and HSE functions define the controls implemented; cybersecurity and audit verify that those controls are effective. The clinical sponsor supplies requirements, data, and acceptance criteria and ultimately receives traceable, patient-specific devices. Logistics interfaces at the boundaries of distribution level.

2.1.1.4 Pre-conditions

We consider the following prerequisites essential for a workable, in-condition case study. First, we maintained multiprotocol wireless coverage across the lab—private 5G, Wi-Fi, BLE, and LoRaWAN—under agreed SLAs and with documented site acceptance. Second, we operated an on-premises data center with virtualized storage and compute, complemented by edge nodes to support low-latency analytics. Third, we enforced secure device onboarding (eSIM/eUICC and certificate management), maintain an accurate asset inventory, and implement zero-trust network segmentation.

Fourth, we integrated MES/MOM for command–acknowledgement exchanges and, where motion control requires it, we interface via vendor-specific controllers or TSN gateways. Finally, we completed safety assessments for collaborative robots and provide worker training with appropriate HMIs and wearables.

2.1.1.5 Triggers

We initiated the workflow under four primary conditions:

- a new product introduction or a reconfiguration event (including cell relocation or tool change);

- the scheduling of a production job in which ML inference indicates early failure risk, evidenced by threshold breaches in vibration or temperature streams;

- a request to instantiate a dedicated network slice for latency-sensitive tasks or the detection of SLA degradation;

- and the occurrence of a security incident or a data-quality anomaly such as schema drift.

2.1.1.6 Normal Flow

A medical-grade **Figure 3** 3D printer configured for clinical and device-manufacturing applications was employed. The printer delivers high-resolution fabrication—down to 20 µm—enabling finely detailed geometries and intricate lattice features suitable for patient-specific devices and precise anatomical models for surgical planning. In our case study, we captured the injured finger's geometry with an EinScan 3D scanner and used the scan-derived measurements to design a custom finger splint. The end-to-end workflow comprised digital modeling of the splint, print execution with our ML-assisted monitoring and control, and post-processing that included finishing and final curing. The process is presented in **Figure 3** below:
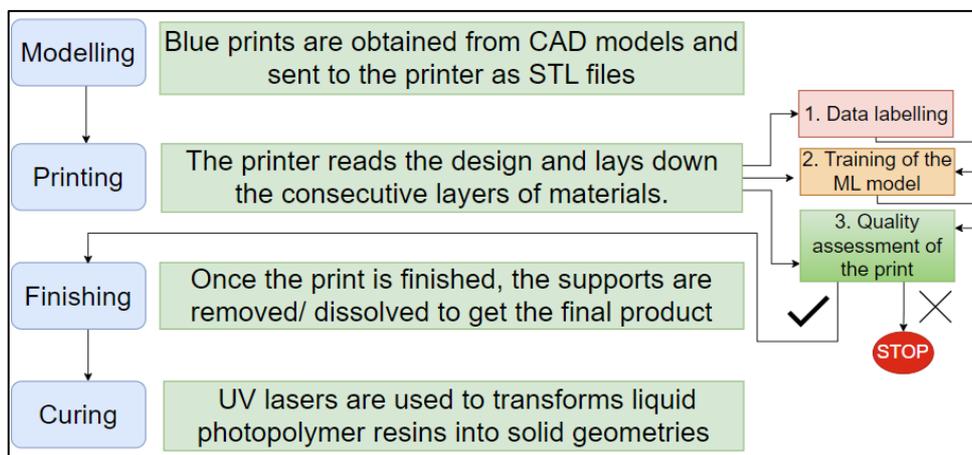
Figure 3 Printing process

A UR3 robotic arm to retrieve parts and transfer them between cells. When the ML system predicted a probable build failure, it sent a halt signal to the printer and, in parallel, instructed the robot to remove the unfinished part and deposit it in the reject bin for later analysis. For successful builds, the robot moved the part to the support-removal cell, where soluble supports were dissolved. After support removal, the ML system signaled the robot to open the curing-station door and load the part for curing. Once curing was complete, the robot transferred the part to the cleaning station and then to the microscope for quality inspection. If the ML system confirmed that the part met patient-specific dimensions and quality thresholds, it instructed the robot to bag the part and place it in the shipping cart.

2.1.1.7 Alternative Flow

In the alternative flow, first address degraded connectivity by failing control traffic over to a redundant channel while buffering non-critical telemetry locally and forwarding it once the link is restored. When a required 5G device class is unavailable, substitute Wi-Fi or wired interfaces for high-bandwidth endpoints and retain LoRa for low-duty metrics, preserving the common data model throughout. If an edge node fails, a hot-standby instance assumes the workload; if no standby is available, place robots in a safe state and continue operations under manual HMI supervision. In the event of a cyber incident, quarantines the affected device, rotate credentials, replay logs to a forensics store, and restore services from immutably logged configurations.

2.1.1.8 Post-conditions

After completion of the use case, the work order is closed in the MES and the full device genealogy is finalized, producing a device history record that links patient scans, process parameters, robot actions, and inspection outcomes to the finished part. All control actions, alerts, and acknowledgements are preserved as tamper-evident logs, and key performance indicators—first-pass yield, scrap, overall equipment effectiveness, and detection/repair times— are computed and published to dashboards and the data lake. If the part passes inspection, conformity to patient-specific dimensions and quality thresholds is confirmed, the item is bagged, and dispatch is scheduled with inventory and shipment records updated. If the part is rejected, it is quarantined, a non-conformance record is opened, and the unit is routed to root-cause analysis; labeled outcomes (pass/fail and failure mode) are appended to the training corpus. Subsequently, the MLOps cycle is triggered: model drift and data-quality checks are reviewed, models are retrained where warranted, new versions are registered, and rollout proceeds under change control. Preventive maintenance is scheduled based on trends in vibration and thermal telemetry. Finally, an auditable close-out pack—logs, KPIs, and configuration fingerprints—is produced for quality, regulatory, and security stakeholders, and lessons learned are captured to refine standard operating procedures and accelerate replication at subsequent sites.

2.1.1.9 High Level Illustration

Figure 4: Overview of MultiProtocol architecture in CONFIRM

Figure 5 overview of Wireless Sensing & Predicting Failures use case

2.1.1.10 Potential Requirements

The portable 5G-enabled medical manufacturing cell must support rapid reconfiguration of printer, UR3 robot, curing/cleaning, and inspection assets with strong device identity, policy assignment, and lifecycle status, backed by a Redis-based data hub that ingests and normalizes telemetry and exposes governed REST/WebSocket APIs to a digital twin, edge analytics, and MES for gated actions such as pause, re-route, or set-point changes.

Radio requirements call for private-5G slicing (URLLC/eMBB/mMTC), engineered indoor coverage and handover with packet duplication or conditional handover, complementary Wi-Fi 6 for non-critical endpoints, LoRaWAN for low-duty telemetry, and disciplined time sync for sensor fusion. Interoperability demands an event schema with versioned registry, accurate time-stamping, immutable audit logs, and secure APIs while safety, security, and privacy require collaborative-robot risk controls and safe states, validated HMIs and interlocks, zero-trust segmentation with mutual TLS and patch governance, SIEM-aligned forensic logging, and GDPR-compliant handling of patient scans and records. Regulatory and quality constraints include alignment to material/process controls for biocompatibility and sterilization chains. The MLOps lifecycle must provide curated labels, drift and data-quality monitoring, reproducible training, staged deployment with rollback, performance SLAs for defect detection, and human-in-the-loop override procedures.

2.1.1.11 Radio Specific requirements

2.1.1.11.1   Radio Coverage

- **Radio cell range (LOS/NLOS).**

Typical indoor NLOS: 30–60 m per small/pico cell in steel-dense spaces.

- o Best-case indoor LOS approx. 100 m.

- o Inside shielded enclosures approx. 3–10 m

- **Public vs. private domain.**

Coverage is constrained to indoor/customer premises (labs, cleanrooms, corridors, staging). No routine coverage of public spaces; optional spillover at loading bays only if required for logistics.

- **Multicell requirement (scope).**

Continuous two-cell overlap is required across all work areas and robot work envelopes to support packet duplication and conditional/make-before-break handover.

- **Handover behavior.**

- o Seamless handover required.
- o Control-path interruption < 20 ms;
- o additional jitter less or equal to 100 us for scheduled control flows.

2.1.1.11.2   Bandwidth requirements
Peak data rate is up to 20 Mb/s (0–20 Mb/s depending on activity); average data rate is 8–12 Mb/s; traffic mode is packet-switched only (IP).

2.1.1.11.3   URLLC requirements
Required latency: < 5 ms one-way; required reliability: > 99.99999%; maximum tolerable jitter must be under 100 us.

2.1.1.11.4   Radio regimens requirements
Desired radio regimen. Private 5G (non-public network) on locally licensed mid-band

2.2   Preliminary 6G use cases

2.2.1 6G-BRICKS (6G- Building Reusable testbed Infrastructures for validating Cloud-to-device breakthrough technologies) Use Cases

6G-BRICKS Principles of Operation

The 6G-BRICKS project (https://6g-bricks.eu/) offers a trusted, agile and evolvable 6G Experimentation Facility, federating two experimentation platforms in Belgium (KUL) and France (EUR) from previous 5G-PPP initiatives under a Core Site (ISI/ATH) acting as the facility entry point, and offering experimentation services. The Facility will be accessible by third party consortiums in the framework of Open Calls, vertical application owners, as well as experimenters from the vertical and component industry. For the detailed list of requirements and description of test beds used in the Use Cases described in the next sections refer to the project deliverable D2.3[2]. These use cases guided the design of 6G-BRICKS overall architecture, containing:

1. **Management Plane Disaggregation**, in line with the "Network of Networks" 6G vision. 6G networks of the future are expected to federate 6G RANs and Cloud-Edge-Continuum deployments from multiple stakeholders, hence will add a level of unprecedented management complexity due to the sheer scale and heterogeneity of the orchestration domains involved. Decentralized orchestration and analytics collection paradigms need to be considered, to increase the scalability of 5G Network data analytics function (NWDAF) and Management Data Analytics Service (MDAS) functions.

2. **AI Explainability**, to be delivered in the application domain of Compute and Network infrastructures. The output of the Explainable AI (XAI) model, along with the Machine Learning (ML's) prediction, is communicated to the experimenters and domain manager. XAI outputs can be attributive, to investigate why the model produced an output and contrastive, to explain why the model did not give another outcome that is different from the current output.

3. **Intent-driven Human-in-the-loop experimentation**. A web-based experimentation plane should act as the entry point to the Facility; mature frontend elements and experimentation engines will be leveraged and extended as part of the Experimentation Plane, to deliver intent-driven, human-in-the-loop experimentation workflow. This allows automated onboarding and instantiation of vertical applications, while providing the actionable explanations that provides an answer to how they can achieve a more desirable outcome (i.e., performance optimization).

4. **Cognitive Compute Continuum.** The 6G design envisions close integration of computation and communication, deploying a new type of network architecture for (extreme) edge to cloud continuum resource orchestration by interconnecting computing and storage resources located anywhere in the network. Cognitive services must be adopted, conducting complex placement and scaling operations, and delivering intelligence on demand (when/where needed) across the communication system.

5. **6G RAN technologies,** including support for RIS and cell-free networking, aiming to support the increased requirements of emerging applications, like the Metaverse and Digital Twinning. These emerging applications exceed the current and future capabilities of 5G networks, both in terms of Key performance Indicators (KPIs) that must be supported, and in terms of their requirements on an ultra-dense computational infrastructure.

6. **Reusable, self-contained enablers with Open APIs ("Bricks")**. Towards this scope, O-RAN interfaces are adopted to ensure the openness and reusability of the developed components, while supporting on-device experimentations via xApps.

---

[2] Deliverable D3.2 "Requirements and Specifications Analysis" https://6g-bricks.eu/wp-content/uploads/2024/07/D2.3_Requirements-and-Specifications-Analysis_v1.0.pdf

20

Summary of 6G-BRICKS Use Cases

The 6G-BRICKS project aims to deliver an evolvable 6G experimentation facility that will federate two testbeds from KUL and EUR, under a common set of experimentation tools deployed at the ISI/ATH site. Accessibility by external entities and openness are key objectives of 6G-BRICKS testbeds, where breakthrough components and technologies are built around the concept of reusable and interoperable "bricks". These are typically based on Beyond-5G technologies that started in previous research initiatives (e.g., in the framework of the 5GPP ICT-52 call) ensuring already existing infrastructures. In what follows the 6G-BRICKS testbeds and their planned interconnectivity in an end-to-end 6G facility is detailed.

The 6G-BRICKS facility will securely interconnect the 6G sites from KUL and EUR with the Experimentation Plane, offered by ISI/ATH, which provides access to experimenters. High bandwidth connectivity will be ensured between the experimentation servers deployed at ISI with the KUL and EUR Sites, leveraging on the fast GEANT VPN network, while security perimeters will be setup via SDP gateways deployed at the 6G Sites. To ensure the security of the facility, Virtual Private Network (VPN) tunnels will be established in a programmatic manner, using a VPN as a Service (VPNaaS) Security Service. This is required within the project to set up a secure channel where a new, unauthenticated experimenter connects to a testbed, aiming to access its features via some exposed APIs (e.g. attempting to connect to some ORAN APIs). In such scenario, the user's request will be derived towards the Security Orchestrator (SO) and the SDP framework over encrypted connections, so that the SDPC can setup the adequate isolated perimeters and secure channels between the experimenter and accessed resources (I.e. services offered by the testbed). In the end, the user gains access and visibility towards the infrastructure. Such a service shall be based on an open-source solution, such as OpenVPN or Wireguard; and potentially be configurable at runtime for further tailoring. Like the decision of the underlying L2 or L3 network, the decision criteria between the available open-source solutions will be guided by factors like throughput, reliability or flexibility. Two Use-Cases are addressed in 6G-BRICKS and are described in the next sections.

2.2.1.1 Use Case 1: Metaverse as an enabler of a Modern Workplace

2.2.1.1.1 Description

Metaverse requires significant KPI improvements by an order of magnitude beyond what is available in 5G systems, while social interactions are supported via the latest advances in XR/VR technologies. The Metaverse leverages XR/VR technologies to enable social interactions in virtual spaces, aligning with the ongoing digitalization of societies and remote collaboration. Videoconferencing is a key aspect of digital transformation, and disruptive technologies can enhance it by mitigating social interaction limitations. Untethered VR already pushes the limits of 5G network capabilities, requiring high capacity and low latency to achieve impressiveness. The Metaverse takes this further, enabling users to interact in 360 XR environments, and recent advancements in volumetric video capture allow multiple users to be co-present in 3D space. The network densification via distributed cell-free can make untethered Metaverse UCs a reality, enabling immersive social interactions with acceptable quality of experience. The Multi-point Control Unit (MCU) element handles the real-time processing of holograms, ensuring synchronization and streaming a 360 VR Sphere of the VE to each participant, receiving and transmitting multiple MPEG 3D video streams with volumetric videos.

2.2.1.1.2 Source

The 6G-BRICKS European SNS JU Project – https://cordis.europa.eu/project/id/101096954

2.2.1.1.3 Roles and Actors

Two separate Proof of Concepts (PoCs) scenarios are outlined below with the corresponding actors:

- PoC1: Holo-conferencing in a Virtual Meeting room: a holo-conferencing scenario is demonstrated in this PoC leveraging ultrahigh-speed with low-latency communications (uHSLLC) achievable only with distributed cell-free technologies. It seeks to implement and integrate a holographic conferencing system in a broadcast environment, integrating the holo-conferencing techniques into a virtual production pipeline. The holographic conferencing system will allow different remote participants to engage in a conversation or discussion while being visualized in holographic form within a virtual environment through VR headsets. At the same time, thanks to the integration of the systems, the audience will be able to view the holographic discussion presented by a host within the TV program. The participating actors for the realisation of this PoC are KUL, NI, ISRD, BRAIN and EKTA.

- PoC2: Virtual Team Building Activities: according to this scenario, Virtual Team Building activities demonstrate ultrahigh data density (uHDD) communications and Joint communications and sensing enablers. In this PoC the VR environment is divided into two separate rooms / dependences, each one including a set of interactive and collaborative activities to be performed in teams. The contributing actors for the implementations of this PoC are KUL, NI, ISRD and I2CAT.

2.2.1.1.4 Pre-conditions

KUL operates a distributed massive MIMO testbed, and a private 5G Standalone Network for end-to-end experiments. The distributed testbed has been extended to allow more distributed and cooperative processing (the controllers were split, and fibres were added for distribution), mmWave connectivity and other advanced features such as full duplex communication, joint communication and sensing, as well as public access through Fed4FIRE. In addition to remote access of the testbed, the KU Leuven has also provided multiple publicly available datasets. KUL also has a full antenna measurement facility, a network analyser (67 GHz PNA-X) and a 4-channel 110 GHz real-time oscilloscope. For processing, the group also owns several GPU servers with multiple state-of-the-art GPU and CPUs.

2.2.1.1.5 Triggers

Initiation of a multi-user XR session requiring real-time holographic communication. Detection of network congestion or user request for improved QoS/QoE also triggers rate adaptation or resource reassignment via APIs. User requests for initiating multi-party XR sessions or detection of conditions requiring bandwidth adaptation or offloading (e.g., network congestion, user join/leave events, or location-triggered events).

2.2.1.1.6 Normal Flow

1. XR application (HoloMIT) requests resource discovery using CAMARA Edge APIs.

2. SOP's Open Exposure Gateway queries local and federated zones via Federation Manager.

3. Application metadata and artefacts are onboarded through the SOP to selected zones.

4. Application instantiation proceeds via SRM and orchestration triggers deployment on edge nodes.

5. NaaS APIs allow runtime adaptations e.g. bandwidth prioritisation, QoS provision, or ATSSS activation.

6. Performance metrics and telemetry are gathered via SOP interfaces and exposed to the AP.

2.2.1.1.7 Alternative Flow

If the primary testbed does not have sufficient resources, the Federation Manager will redirect the deployment request to a federated SOP. If the federation fails, fallback to degraded XR service using pre-cached low-fidelity assets is applied. Rate limitation and adaptive encoding is performed dynamically to maintain user QoE.

2.2.1.1.8 Post-conditions

Application session ends and teardown commands are issued via SOP. Resources are de-allocated, performance logs stored, and any gathered metrics can be published for research or SLA evaluation. Successful or partial completion of XR sessions. KPIs, usage logs, and QoE metrics are stored. Resources are released and the federation registry is updated. Optional dataset anonymization and publication may follow for standardization or research.

2.2.1.1.9 High Level Illustration

The **Figure 6** provides a high-level architecture diagram for this envisioned platform and PoC. Its main building blocks are briefly introduced next:

- **Single-sensor Volumetric Video Capture and Reconstruction Module**: it is responsible for capturing a frontal viewpoint of a human body (i.e., hologram) in real-time, by using an off-the-shelf RGB-D sensor (e.g., Azure Kinect). The captured RGB-D stream is then processed / encoded for its distribution over the network to the remote participants. The captured stream is also sent to the local rendering engine to provide a self-representation of the local participant.

- **Live Presenter Module**: it is responsible for capturing, encoding and delivering the silhouette of the show presenter from e.g., a Chrome key room.

- **Media Server(s)**: cloud server(s) that store and provide different media assets to be used for / in the Metaverse experience (e.g., the VR environment, videos to be presented, interactive 3D elements).

- **Orchestrator Module**: It provides service and session management features to enable the desired multiuser experiences, and it additionally interfaces the required networking and in-cloud media processing elements.

- **Packet Reflector Module**: it oversees forwarding incoming streams to the target destinations within the same shared session.

- **Remote Rendering Module**: it oversees providing interactive 2D or 360º video streams from the 3D VR experience to lightweight client players (spectators).
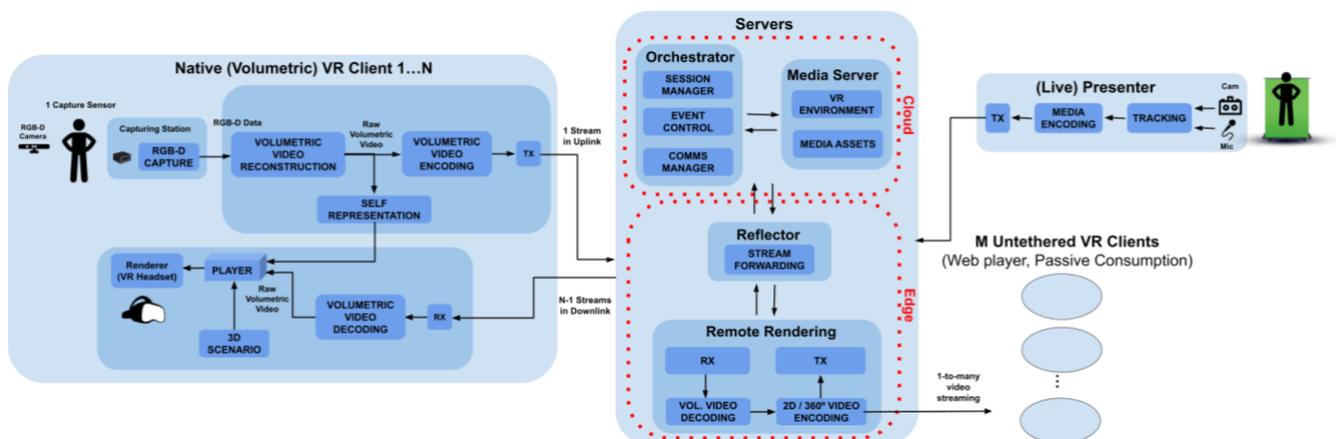
Figure 6: High-level system architecture of the envisioned platform and associated PoC 1

2.2.1.1.10 Potential Requirements for PoC1: Holo-conferencing in a Virtual Meeting room

The technical and user requirements for the system to be implemented are presented below. An agreed-upon set of requirements contributes to a design that meets the expectations of stakeholders and reduces discrepancies that may occur during system verification, testing, and validation.

- Required minimum latency

When working with a real-time communication system for broadcast, the goal is to achieve minimal latency, ideally less than 1 second.

- Visual Quality

  The proposed system requires a high minimum visual quality for proper integration and real-time result previewing.

- Required Feedback in Communications

  Video and audio feedback is necessary in the system to control the interaction between the roundtable and the presenters.

- Scene control

  Scene and 3D elements control will be carried out by a system operator who assists the teachers in managing the entry and exit of interactive elements as well as the integration of remote presenters.

- Image Resolution

  The final resulting image must have a high-quality HD resolution to showcase the integration of the elements in the virtual scenario.

- Required Power

  The power of the graphic stations and the image quality must allow for high-quality real-time chroma keying of the presenter.

Other requirements for PoC1: Holo-conferencing in a Virtual Meeting room

Requirements for the holographic comms sub-platform:

- Support for at least 4 holo-ported users in a shared session

- End-to-end latency between holo-ported users <=150ms

- Holograms with a resolution of >=100K voxels (i.e. pixels in 3D) per frame and >=15fps (which requires inter-frame arrival periods, and this jitter, <=66ms)

- Here are the user requirements, as seen from the audience's point of view, which are essential for achieving a high-quality experience. The audience will access the experience through a broadcast system, so it is essential to ensure maximum visual quality.

Optimal Visual Quality	The visual quality of the results must be as high as possible to create a good quality experience.
High-Quality Sound	The feedback in the holographic system communication largely depends on audio transmission. For this reason, audio must be transmitted and received with the highest possible quality.
Integration in the Environment	A realistic integration into the environment is sought using virtual shadows, lighting, and colour adjustments to make the presenter and the roundtable appear fully integrated within the virtual scene.

2.2.1.1.11 Potential Requirements for PoC 2 Multiuser Team Building Activities with Advanced Full Volumetric Capture and In-cloud Media Processing

Relevant Requirements for viewers/participants quality of experience

The following are the essential requirements for the proposed system.

Table 1: Requirement U1P2-RT1.

ID	U1P2-RT1	Type	Latency
Name	Required minimum latency		
Description	The communication latency between collaborating users shall be lower than 150ms, while the inter-group switching delay (i.e. a team member switching to another team in the collaborative session) shall be lower than 0.5s.		

Table 2: Requirement U1P2-RT2.

ID	U1P2-RT2	Type	Quality
Name	Visual Quality		
Description	The proposed system requires a high minimum visual quality for 3D hologram reconstruction and perception from 6 Degrees of Freedom (DoF) environments.		

Table 3: Requirement U1P2-RT3.

ID	U1P2-RT3	Type	Scalability
Name	Scalability		
Description	At least 2 different teams with at least 4 members each shall be supported.		

Table 4: Requirement U1P2-RT4.

ID	U1P2-RT4	Type	Interaction
Name	Interaction Features		
Description	Support for >=4 multi-modal interaction features (e.g., shared media presentation, shared interactive boards, collaborative assembling of 3D pieces, manipulation of 3D elements).		

Table 5: Requirement U1P2-RT5.

ID	U1P2-RT5	Type	Interaction
Name	Free Tele-porting		
Description	Free tele-porting within the VR environment, containing two separate rooms.		

Table 6: Requirement U1P2-RT6.

ID	U1P2-RT6	Type	Processing
Name	Processing Resources		
Description	The involved cloud servers and clients shall support the real-time processing and rendering of clustered scenes (for each team, room), each with high resolution holograms (volumetric video streams).		

As for user requirements, the same ones for PoC1 apply, and an additional one is elicited.

Table 7: Requirement U1P2-RU1.

ID	U1P2-RU1	Type	Interaction
Name	Natural and Effective Interaction		
Description	The envisioned multi-modal and multi-user interaction features shall be effective (i.e., precise and with minimum latency) and shall be perceived as natural as possible (i.e., as closer as possible to real-time interactions.		

KPIs

Requirements for the holographic comms sub-platform:

- Support for at least 6 holo-ported users in a shared session, and at least 2 separate teams,

- End-to-end latency between holo-ported users <=250ms when using cloud processing,

- Holograms with a resolution of >=100K voxels per frame and >=15fps,

- XR-MCU being able to independently handle two subsets of streams (2 teams),

- Interaction delays <=100ms

2.2.1.2 Use Case 2: 6G applications for Industry 4.0

This Use Case focuses on the applications of Industry 4.0 and how 6G technology can enhance the efficiency of operators using autonomous robots, digital twinning, and XR. The market for autonomous driving and DT in industrial applications is growing rapidly, but its impact depends on the collection of high-volume data with low latency, which 6G networks can provide. To meet the requirements of precise modelling, simulation, and verification, a Tbps-level transmission rate is often necessary. Specifically, a key to this transformation are autonomous shuttles robotic systems designed to move goods from one location to another. They use sensors, cameras, and CPUs to navigate around facilities, optimizing productivity and space use. All the software and technologies within shuttle systems need to be part of one cohesive system. This system needs a mechanism for wireless M2M communication and dependable management systems, enabling remote monitoring and control.

Security and privacy are also crucial considerations, and 6G networks can support data storage, collection, processing, training, and modelling in both centralized and distributed architectures. Additionally, Next Generation Mobile Networks (NGMN) recognize autonomous robots and enhanced Machine Communication as key 6G use-cases, which require extreme low-latency communications and high reliability to synchronize robots among themselves and with remote controllers.

2.2.1.2.1 PoC 1: Autonomous robots in Industry 4.0

In the realm of Industry 4.0, the integration of autonomous robots has become imperative for achieving efficient and scalable manufacturing processes. However, the challenges of communication in signal-obstructed factory environments have persisted. In response, Reconfigurable Intelligent Surfaces (RIS) are emerging as a promising solution to enhance connectivity. To further address these limitations, the introduction of 6G technology promises lower latency and higher capacity, laying the foundation for seamless communication. This use case showcases the synergy between a 6G testbed and RIS systems, demonstrating how they enable uninterrupted robotic autonomy and mission execution, even in non-line-of-sight scenarios within industrial settings. The methodology outlined herein encompasses a structured approach, including environment setup, parameter optimization, simulated obstacle avoidance using reinforcement learning, and performance evaluation on physical topologies. Through the development and testing of an autonomous driving system for the TurtleBot4 mobile rover, leveraging ROS, machine learning algorithms, and reinforcement learning, this research aims to advance TurtleBot4's autonomous capabilities efficiently and effectively.

Scenario

In the context of Industry 4.0 automation, the demand for efficient and scalable manufacturing processes has led to increased use of autonomous robots. However, these robots face communication challenges, especially in signal-obstructed factory environments. Reconfigurable Intelligent Surfaces (RIS) are being explored as a solution to enhance connectivity. To address existing network limitations, 6G technology is introduced, offering lower latency and higher capacity. This use case demonstrates how a 6G testbed, in conjunction with RIS systems, can ensure seamless robotic autonomy and mission execution, even in non-line-of-sight situations, within industrial settings, all with a focus on mmWave bands and advanced networking technologies. The methodology for this use case encompasses the following stages:

- Setting up the project environment, which includes installing Linux, the ROS, and the necessary TurtleBot packages.

- Testing and fine-tuning the SLAM and navigation parameters to optimize TurtleBot4's basic functionality.

- Conducting autonomous driving tests on and implementing and testing reinforcement learning algorithms for obstacle avoidance in a simulated environment.

- Evaluating the TurtleBot4's autonomous driving performance on a physical topology.

The scope of this use case is to develop and test an autonomous driving system for the TurtleBot4 mobile rover using ROS, machine learning algorithms, and reinforcement learning in industrial environments. The research includes designing and constructing a custom industrial environment for testing the autonomous driving system and evaluating the system's performance in actual and simulated environments.

The TurtleBot 4 Lite and TurtleBot 4 Standard are both mobile robot platforms based on the iRobot® Create® 3 base, sharing many features such as size, wheel diameter, ground clearance, the on-board computer (Raspberry Pi 4B 4GB), maximum linear and angular velocities, maximum payload, operation and charging times, and the type of Lidar used. However, there are significant differences between the two models that can influence a researcher's decision (**Figure 7**):

1. **Camera**: The TurtleBot 4 Standard is equipped with an OAK-D-Pro camera, which is superior to the OAK-D-Lite camera in the TurtleBot 4 Lite. The Pro version typically offers better performance in terms of image quality and depth sensing, which is crucial for more complex navigation and object recognition tasks.

2. **User Power and USB Expansion**: The TurtleBot 4 Standard provides more options for power outputs and USB connectivity. This includes higher current outputs and additional USB 3.0 Type C ports, which are beneficial for connecting more or higher-powered external devices and sensors.

3. **Status LEDs and Buttons**: The TurtleBot 4 Standard features additional status LEDs and user buttons, which enhance the user interface for direct interaction and provide immediate visual feedback on the robot's status. This is useful for debugging and easier control of the robot.

4. **Bluetooth Controller**: Only the TurtleBot 4 Standard includes the TurtleBot 4 Controller, which can be a crucial element for easier manual control and for tasks requiring precise manoeuvring.

Figure 7: Turtlebot4 Components

Given these differences, we ended up choosing the TurtleBot 4 Standard over the TurtleBot 4 Lite for two important reasons:

1. **Enhanced Sensor and Expansion Capabilities**: The superior camera, additional power outputs, and extensive USB expansion support in the TurtleBot 4 Standard allow for more complex projects and experiments. These features make the Standard model a more versatile and future-proof choice, enabling a wider range of applications, from advanced navigation and mapping to sophisticated sensor integration and data collection.

2. **Improved User Interface and Control**: The inclusion of a Bluetooth controller, additional user buttons, and comprehensive status LEDs make the TurtleBot 4 Standard more user-friendly and accessible. These features simplify the process of controlling the robot and interpreting its status, which is especially beneficial for educational purposes and rapid development cycles.

Although the TurtleBot 4 Lite might be a more cost-effective and lighter option, the advanced features and greater expandability of the TurtleBot 4 Standard offer better value for our specific needs, particularly in terms of enhanced operational capabilities and ease of use.

This methodology guides the development and testing of TurtleBot4's autonomous driving. It covers environment setup, parameter optimization, simulated obstacle avoidance using reinforcement learning, and performance evaluation on a physical topology. This structured approach is key to advancing TurtleBot3's autonomous capabilities efficiently.

Building blocks and Platform

In the context of automated manufacturing solutions, the architecture in **Figure 8** involves the integration of various components to enhance the efficiency, scalability, and communication in factory operations.
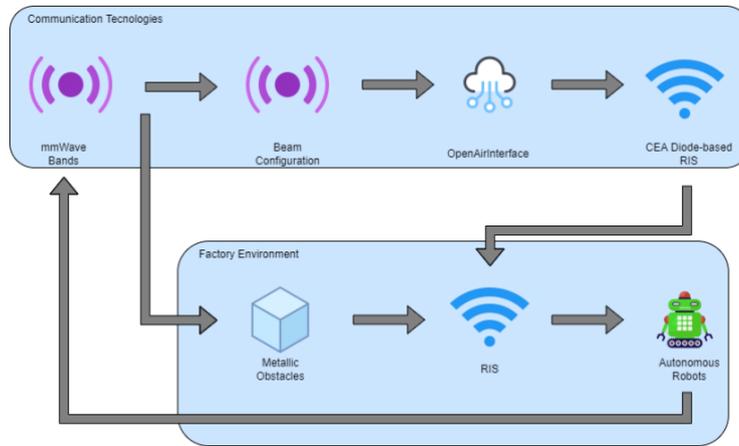
Figure 8: High Level Architecture (TurtleBot)

This architecture primarily focuses on addressing the challenges posed by metallic obstacles within the factory, the use of mmWaves for communication, the utilization of Reconfigurable Intelligent Surfaces (RIS), and the requirements for autonomous robots and control centres.

- **mmWave**: mmWave technology uses high-frequency millimetre waves for wireless communication, offering high data capacity. It's used for connecting autonomous robots in manufacturing but can be obstructed by obstacles.

- **Beam Configuration**: Beam configuration involves directing millimetre wave signals with precision to avoid obstacles, ensuring reliable and uninterrupted communication in manufacturing environments.

- **Open Air Interface**: is a software-defined radio access network platform that configures and controls communication infrastructure, playing a key role in managing mmWave signals and other network elements.

- **CEA Diode-Based RIS for Metallic Obstacles**: CEA diode-based Reconfigurable Intelligent Surfaces (RIS) are designed to manipulate signals around metallic structures, reducing signal blockage and enhancing communication reliability in manufacturing facilities.

- **RIS (Reconfigurable Intelligent Surfaces)**: RIS are specialized surfaces that control the direction of transmitted signals, reflecting them in chosen directions to ensure signal paths are maintained, especially in complex environments with obstacles.

- **Autonomous Robots**: Autonomous robots are employed in manufacturing to perform complex and repetitive tasks efficiently. They rely on seamless network communication to receive instructions and transmit data for effective operation and coordination.

- **6G and 5G Networks**: The architecture considers the use of advanced telecommunications networks, such as 6G and 5G. While 5G networks are employed to expand the scale of connected devices in manufacturing operations, 6G networks offer lower latency, allowing for more precise and efficient control of the process. The lower latency of 6G is particularly important for autonomous robotics applications where real-time control is critical.

Suggested platforms/tools:

1. **ROS (Robot Operating System):** It is an open-source middleware platform for developing autonomous robots. ROS provides a framework for creating and running robot applications across multiple hardware platforms.

2. **Robot Operating System Industrial (ROS-I):** This is an extension of ROS specifically designed for industrial automation.

3. **MATLAB Robotics System Toolbox:** Designing, simulating, and testing robotics applications. It supports various robot manipulators, mobile robots, and unmanned aerial vehicles (UAVs).

4. **Microsoft Robotics Developer Studio:** Developing robotic applications using Microsoft technologies. It supports various programming languages, including C++, C#, and Visual Basic.

5. **Intel RealSense SDK:** This is a software development kit for integrating Intel RealSense cameras into robotics applications. It provides APIs for various functions, including depth sensing, 3D scanning, and facial recognition.

This architecture combines mmWave communication, beam configuration, RIS technology, Open Air Interface, CEA diode-based RIS for metallic obstacles, autonomous robots, and control centres to create a robust and efficient system for automated manufacturing. It addresses the challenges posed by signal blockage and the need for low latency, ultimately enhancing the scalability and performance of manufacturing operations.

2.2.1.2.1.1 Potential Requirements for PoC 1: Autonomous robots in Industry 4.0 (viewers/participants quality of experience)

The following are the essential requirements for the proposed system.

Table 8: Requirement U2P1-RT1.

ID	U2P1-RT1	Type	Technical
Name	Low Latency Communication		
Description	The system must provide low-latency communication, with minimal latencies ranging from 0.1 to 20 milliseconds, to enable precise and efficient control of autonomous robotic operations within the industrial environment.		

Table 9: Requirement U2P1-RT2.

ID	U2P1-RT2	Type	Technical
Name	Seamless Connectivity with RIS		
Description	The system must seamlessly integrate with Reconfigurable Intelligent Surfaces (RIS) to overcome signal obstructions and ensure uninterrupted connectivity for autonomous robotic systems, even in non-line-of-sight (NLOS) situations.		

The tables presented below outline the user requirements for the UC2 and more specific about POC1.

Table 10: Requirement U2P1-RU1.

ID	U2P1-RU1	Type	Operational
Name	Efficient Robot Mission Management		
Description	Users require an efficient robot mission management system that allows translation of operator commands into missions, facilitating agile execution in indoor industrial environments using a 6G network.		

Table 11: Requirement U2P1-RU2.

ID	U2P1-RU2	Type	Operational
Name	Scalability and Uninterrupted Connectivity		
Description	Users demand a scalable solution that maintains uninterrupted connectivity, even when autonomous robot systems go out of the line-of-sight (LOS), enabling the expansion of manufacturing operations and the control of a growing number of connected devices.		

KPIs

Requirements for autonomous robots in industry 4.0:

- Latency: measure the system's average latency in milliseconds, aiming to maintain it within the 0.1-20ms range to ensure precise control of robotic operations.

- Connectivity Reliability: track the percentage of time the robotic system remains connected and operational, even in NLOS situations with the help of RIS (approximately 99%).

- Mission Execution Efficiency: evaluate the efficiency of the system in executing missions, monitoring the percentage of missions completed successfully within expected time frame (approximately 95%).

2.2.1.2.2 PoC 2: AR inspection of Industry 4.0 digital twin on site

To handle the limitations of visual inspection as the primary technique for identifying malfunctions in industries where complicate machinery and/or heavy machinery exists, leading to risks for accidents and downtime, this PoC proposes the use of digital representation to enhance critical information availability for indoor scenarios. This digital representation occurs using augmented reality (AR) to provide a digital twin situation enabling a mirror-like experience for remote and local operators. AR technology overlays digital information onto the real world, seamlessly blending virtual elements with the user's physical environment. Experts from the manufactures of the equipment or trained personnel can reside far from the industrial setting and can support multiple sites. Field personnel with the appropriate gear (tablet and AR Glasses) can communicate with the expert and be provided with guidance. The expert is there with the field personnel understanding and determining the best course of action with accessible libraries of information not available to field personnel. The expert can provide specific guidance eliminating the need to revisit the problem again or maintenance issue again.

Scenario

The scenario envisaged as a PoC involves the remote inspection of a hardware component in the factory as it is part of the factory's regular inspections. Prior to the scenarios execution a scan is performed on the hardware component, so that the system becomes familiar with it. Additionally visual markers are placed on the object to denote maintenance and control action points (**Figure 9**).

During a scheduled maintenance visit the field personnel adorn the equipment described below and engage in a digital twin environment with the expert sitting far away. Using the media available, video and audio, the expert guides the field personnel in performing the inspection.
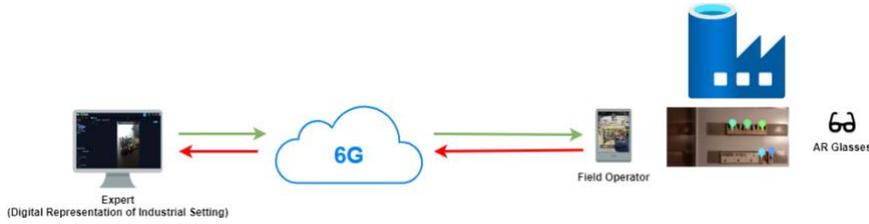
Figure 9: Conceptual Architecture

<u>**Building blocks and Platform**</u>

The following are the major components of the system:

1. **EPSON AR Glasses**: Smart glasses for use in indoor and outdoor. The wearer can visualize and perform actions based on guidance provided by the remote expert.

2. **Digital Twin**: An enhanced Content Management System (CMS) which allows the expert to make use of the digital twin of the industrial equipment enabling the guidance to the field personnel. It is a fully-fledged management system with the following modules:

    a. Digital Twin Environment

        i. Audio

        ii. Video

        iii. Text

        iv. Schematic Diagrams

    b. Document Library

    c. User and Role Management

    d. Multi-lingual user interface

    e. Maintenance and Inspections

3. **Tablet device**: Provides the main user-interface device for the AR platform enabling the communication of the AR glasses with the remote expert device. Contains wireless features enabling the bi-directional communication with the remote expert. It also acts as the training device for the system enabling all the industrial equipment to be captured and catalogued.

2.2.1.2.2.1 Potential Requirements for PoC 2: AR inspection of Industry 4.0 digital twin on site (for viewers/participants quality of experience)

The following are the essential requirements for the proposed system.

Table 12: Requirement U2P2-RT1.

ID	U2P2-RT1	Type	Latency
Name	Required minimum latency		
Description	When working within the digital twin environment by the expert the goal is to achieve minimal latency, ideally less than 2 seconds.		

Table 13: Requirement U2P2-RT2.

ID	U2P2-RT2	Type	Quality
Name	Visual Quality		
Description	The proposed system requires a high minimum visual quality for proper integration and real-time result previewing.		

Table 14: Requirement U2P2-RT3.

ID	U2P2-RT3	Type	Feedback
Name	Required Feedback in Communications		
Description	Video and audio feedback is necessary in the system to control the interaction between the expert and the field personnel.		

The tables presented below outline the user requirements for the UC2 and more specific about the POC2.

Table 15: Requirement U2P2-RU1.

ID	U2P2-RU1	Type	Quality
Name	Optimal Visual Quality		
Description	The visual quality of the final results must be as high as possible to create a good quality experience.		

Table 16: Requirement U2P2-RU2.

ID	U2P2-RU2	Type	Sound
Name	Uninterruptible Audio		
Description	The audio between the factory and the expert should be uninterrupted to ensure a successful remote inspection.		

KPIs

Requirements for AR inspection of Industry 4.0 digital twin on-site:

- Real-time Data Delivery: Measure the system's ability to deliver real-time data to managers and supervisors for analysing component functions (data refresh rate of a minimum 10 frames per second to provide a smooth visual experience) aiming for minimal latency (less than 2 seconds)

- Accident Reduction: Track the percentage reduction in accidents and downtime attributed to the use of DT and AR in identifying malfunctions (at least 20-30% reduction in accidents and 25-40% in downtime caused by equipment malfunctions).

- Platform Compatibility: Evaluate the system's compatibility with multiple platforms and tools, assessing its adaptability and ease of integration with existing infrastructure (95% success rate when integrating the system with various industrial platforms and tools).

Radio Specific requirements

The following envisaged KPIs have been anticipated:

- *Extreme data rates*: For both indoor and outdoor connectivity, peak data rates up to 1 Tbps are envisioned. However, particularly for the user-experienced data rate, it is guaranteed to reach 1Gbps for 95 % of the user locations.

- *Enhanced spectral efficiency and coverage*: The peak spectral efficiency can be increased up to 60 b/s/Hz by means of improved MIMO and modulation schemes. Again, the *user experienced spectral efficiency* is envisaged to reach 3 b/s/Hz. This can be achieved through new PHY layer techniques to allow for broadband connectivity in high mobility scenarios.

- *Very wide bandwidths*: Bandwidths up to 10 GHz can be supported in mmWave bands, while up to 100 GHz can be reached in THz and visible light bands.

- *Enhanced energy efficiency*: Since sustainability is one of the key issues on 6G technologies development, it is expected to pay special attention in achieving better energy efficiency. For the case of transmission energy, the consumption should reach up to 1 terabit per Joule.

- *Ultra-low latency*: Latency shall be brought down to 0.1ms for bandwidths of the order of 10 GHz. The latency variations (jitter) should reach down to the extreme limit of 1 µs.

- *Extremely high reliability*: Some new use cases require extremely high reliability up to $1-10^{-9}$ to enable mission and safety-critical applications.

2.2.2 6G-DALI (6G DAta and ML operations automation via an end-to-end AI framework) Use Cases

Overview of 6G-DALI Architecture

The 6G-DALI framework (https://6gdali.eu/) delivers a sophisticated orchestration and experimentation platform for next-generation AI/ML-based networking in 6G environments. Overall, 6G-DALI provides a robust foundation for AI-native 6G innovation by integrating MLOps/DataOps, digital twin technology, and decentralized orchestration into a seamless research and experimentation framework. The architecture is modular, secure, and aligned with European data sovereignty principles, enabling trusted experimentation across academic, industrial, and vertical domains. The high-level architecture of 6G-DALI, depicted in **Figure 10,** outlines the main components of the end-to-end AI framework. It is designed around two main pillars:

1. AI experimentation as a service via MLOps

2. Data and analytics collection and storage via DataOps

Both above support the automation of MLOps and DataOps workflows, respectively. The framework serves both ML developers, focused on optimizing 6G infrastructure and enabling closed-loop, zero touch management, and vertical users developing ML models to enhance their 6G services. The framework also interfaces with a range of 6G-testbeds, encompassing RAN, core networks (CN), and computing infrastructure (from edge to cloud), with some also including extreme-edge devices. Moreover, there is at least one Digital Twin (DT) for large-scale testing. Central to this architecture is the creation of a 6G Dataspace that links data producers and consumers, enabling dataset generation, referencing, and sharing. Relied on the IDS and Gaia-X principles, this Dataspace promotes secure, trusted, and sovereign data exchange among stakeholders (ML developers, verticals, operators, and vendors) paving the way for a robust 6G data economy.
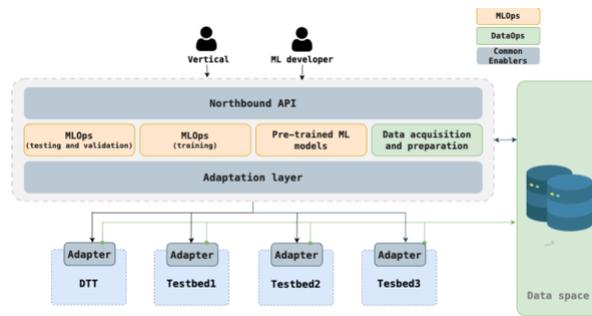
Figure 10: 6G-DALI AI framework - high-level architecture

2.2.2.1 Use Case 1: Data Management and Experiment-on-Demand

The AI framework integrates LLMs for intent parsing and dataset generation workflows. A key architectural innovation is the dual-path data discovery mechanism: GAIA-X-based catalogue lookup for 'cold' data and orchestrated testbed-driven data generation for 'hot' data. The Extract-Load-Transform (ELT) pipeline includes modules for semantic intent translation, dataset normalization, and augmentation. Integration with the 6G Data Space relies on secure APIs and Gaia-X compliant catalogues, supported by Piveau components.

2.2.2.1.1 Description

The Data Management and Experiment-on-Demand use case demonstrates how the 6G-DALI AI framework can fulfil user requests for 6G datasets through two main paths: retrieving existing datasets from a federated 6G Data Space or triggering new experiments when data is not available. It integrates with Gaia-X service catalogues, interprets user intents via LLMs, and supports ELT pipelines to clean, augment, and load data. This enables researchers to discover or generate on-demand high-quality datasets using real-world testbeds, fostering collaborative AI experimentation.

The aim of this UC/PoC is to showcase the capabilities of the 6G-DALI AI framework to satisfy users' requests for 6G data by building the concept of the 6G Data Space, with the participation of all 3 6G testbeds and the TID cloud. The two pipelines supported by 6G-DALI for dataset search will be demonstrated. The first one aims to demonstrate the case that the data is already in the 6G Data Space and discovered through the AI framework using the concept of Gaia-X service catalogue, while the second one showcases the experiment(s) triggering the case when the data is not existing in the 6G Data Space. Both experiments will demonstrate the ELT pipeline, from extracting the semantics of users to translating it to a service catalogue request, and if the dataset is not available translate the request to experiments to run on one or multiple testbeds connected to the project.

The list of objectives includes:

- Demonstrate the integration of the AI framework with 6G Data Space and Gaia-X service catalogue, authorization, and trust.

- Demonstrate the AI framework's ability to translate user requests to data sets to experiments via LLM technologies.

- Demonstrate the ELT pipeline to populate, transform (data cleaning and augmentation), and load to the data space for data analytic collection and storage.

The development of AI and machine learning models for 6G systems and services requires access to large, high-quality datasets. However, such datasets are often scarce, fragmented, or in need of cleaning and standardization to conform with widely accepted formats. In parallel, emerging initiatives like *GAIA-X* and IDSA are pioneering new approaches to secure and interoperable data sharing through concepts such as **Data Spaces, Catalogues, and Marketplaces**. Sectors like transportation and healthcare have already begun adopting these initiatives, demonstrating their value and practicality.

Bringing these advancements into the communications domain presents a unique opportunity to accelerate the development of robust, efficient, production-ready AI/ML frameworks for 6G. As part of this vision, the **Data Management and Experiment-on-Demand Proof of Concept (PoC)** in 6G-DAILI aims to highlight the benefits of adopting these data-sharing principles. A central capability of the 6G-DAILI AI framework is to respond to user requests for 6G data through an integrated and intelligent data ecosystem.

At the heart of this ecosystem is the **6G Data Space**, which interconnects the project's three 6G testbeds, each contributing unique datasets and experimentation capabilities. The PoC supports two operational workflows:

- **Utilization of Existing Datasets**: Leveraging datasets already available in the 6G Data Space, originating from the testbeds.

- **On-Demand Dataset Generation**: Executing targeted experiments on the testbeds to generate new datasets when no existing dataset meets the experimenter's requirements.
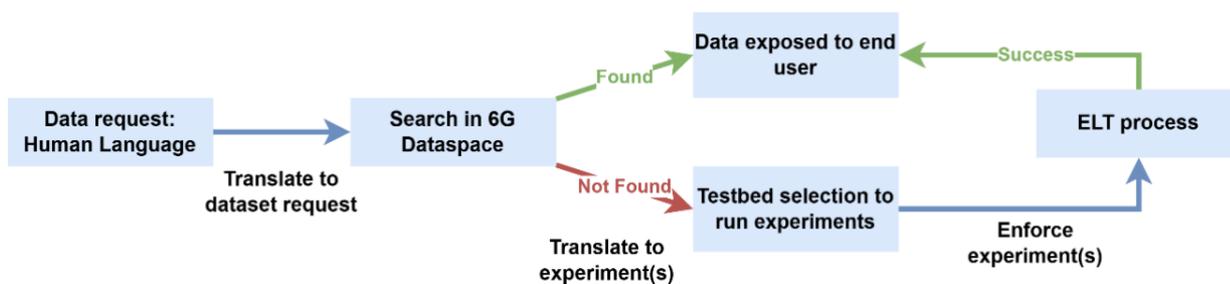
Figure 11: Dataset Discovery process

A key element of this PoC is **dataset discovery**—determining whether a suitable dataset is already available or needs to be generated. To support this, the AI framework will leverage Data Space services developed in alignment with GAIA-X and IDSA specifications, ensuring secure, interoperable discovery and retrieval. Beyond data access, the PoC will also demonstrate how the 6G Data Space enables the creation of **ELT (Extract, Load, Transform)** pipelines. These pipelines will support the transformation of datasets according to user-defined needs, such as data cleaning or augmentation. Once processed, the enriched datasets will be reintegrated into the 6G Data Space, making them available for future reuse by other researchers.

Through these workflows, the PoC will validate the following capabilities:

- **Integration** of the AI framework with the 6G Data Space and GAIA-X services for cataloguing, authorization, and trust management.

- **AI-driven translation** of user data requests into experiments using large language model (LLM) technologies.

- **Execution of ELT pipelines** to populate, transform, and reload datasets into the data space, enabling continuous data enrichment and reuse for analytics and research.

The two experiments to be performed as part of this POC are:

- **Data search and extraction via Gaia-X service catalogue**: This experiment demonstrates the ability of the 6G-DALI framework to fulfil user data requests—also referred to as user intents— using datasets that are already available within the 6G Data Space, without the need to run new experiments. It showcases how the system can intelligently interpret user needs, perform secure and trusted data discovery, and apply automated data transformation procedures.

This experiment validates the framework's ability to translate abstract user intents into precise dataset discovery actions, perform trusted search and access through Gaia-X compliant services, and apply transformation workflows that prepare data for immediate use in AI/ML development or analytical applications. It highlights the seamless interplay between intelligent interfaces, secure data infrastructure, and advanced data processing within the 6G-DALI ecosystem.

- o Experiment begins with a user submitting a data request through the 6G-DALI Northbound API. This interface integrates a large language model (LLM) that interprets user's high-level intent and translates it into a structured search query. The query is then executed against 6G Data Space's service catalogue, which adheres to Gaia-X principles and incorporates tools such as a Piveau Catalogue, with trust/security mechanisms for compliant and reliable data handling.

- o For the purposes of this demonstration, the Data Space is pre-populated with three curated and published datasets originating from the EUR and KUL testbeds. Once the dataset search is complete, the results are presented to the user. If a relevant dataset is found, it can be used immediately or further refined. Upon selection, the system may automatically trigger or allow the user to initiate an ETL (Extract, Transform, Load) pipeline. This pipeline performs data cleaning and augmentation, leveraging algorithms developed within the 6G-DALI project to ensure the dataset is fully tailored to the user's intended use.

- **Data on-demand and enhanced ELT pipelines**: This experiment focuses on demonstrating the complete end-to-end management of a user data request that cannot be satisfied with existing datasets in the 6G Data Space. Unlike scenarios where datasets are readily available through the service catalogue, this experiment activates the second operational workflow of the PoC: the on-demand generation of datasets through targeted experimentation across the project's testbeds. Through this experiment, the PoC demonstrates the seamless integration of intelligent orchestration, AI-powered translation of user needs into experiments, and automated ELT processes within the 6G-DAILI framework. It validates how this framework can dynamically respond to evolving data demands, promote dataset enrichment, contributing to a scalable, collaborative research environment.

- o The process begins when a user submits a data request. Since no matching dataset is found in the Data Space, the request is automatically translated into an experiment description using a dedicated large language model (LLM), designed to interpret and convert high-level data needs into executable experiment plans. Following this translation, a testbed selection mechanism determines the most appropriate infrastructure to run the experiment based on the nature of the request. The experiment is then deployed across the available testbeds: EUR for Open RAN scenarios, ISI for the cloud continuum, and KUL for advanced 6G wireless communication use cases. At least two of these testbeds will be engaged during executions.

- o Once the experiments are completed and raw data is collected, the system triggers an ELT (Extract, Load, Transform) pipeline. This pipeline is responsible for extracting the data from the testbed environments, transforming it through cleaning and augmentation methods developed in 6G-DALI, and finally loading the enriched dataset back into the 6G Data Space. The result is a high-quality, reusable dataset that can be accessed by other users/systems in future interactions.

2.2.2.1.2 Source

- 6G-DALI Project on CORDIS: https://cordis.europa.eu/project/id/101192750
- 6G-DALI Project portal: https://6gdali.eu/
- 6G-DALI on SNS-JU web: https://smart-networks.europa.eu/call-3-stream-b/#6G-DALI

2.2.2.1.3 Roles and Actors

This use case (PoC) is expected to involve the following main actors:

- **Researchers (R)**, stakeholder in academia or industry domains focusing on research and development activities. They need to leverage 6G-DALI's experimentation capabilities to accelerate 6G-related AI innovation, validate theoretical models, and collaborate via open, reproducible tools, environments, and high-quality datasets. **SPW, TUD**

- **Cloud Vendors (CV)**, Cloud service providers offering infrastructure and platforms. In 6G-DALI, they support hosting of ML models, training datasets, and data-sharing environments. **TIF**

- **AI experts and ML developers (AEMD)**, representing the stakeholders that create and optimize the machine learning models through hyperparameter optimization and transfer learning. **IQU**

- **Testbed Business Owner (TBO)**, representing the stakeholders that provide the testbed infrastructure where the experiments are deployed. **ISI**, **KUL**, and **EUR** will provide the testbeds that will be involved in the experiments.

The participants benefit from the PoC in the following ways:

- **Researchers (R) & AI experts and ML developers (AEMD)**, using the 6G-DALI's infrastructure researchers from every domain will greatly benefit from the ability to discover or create on-the-fly datasets that will help them evaluate and test their applications using real-world data in real-world conditions and using real hardware using a unified experimentation environment and automated pipelines.

- **Cloud Vendors (CV)**, demonstrate value as infrastructure providers for scalable, data processing and transformation workloads across cloud-native and edge environments.

- **Testbed Business Owner (TBO)**, validate their infrastructure's capability to share datasets and generate on-demand ones for the 6G domain.

2.2.2.1.4 Pre-conditions

A user has a data need for a specific 6G-related experiment or training scenario. The AI framework and its integration with the Data Space, testbeds, and LLM components must be operational.

2.2.2.1.5 Triggers

A user request for a dataset—either fulfilled by existing entries in the 6G Data Space or requiring new experiments to be executed across testbeds.

2.2.2.1.6 Normal Flow

1. User submits a data request via intent-based Northbound API.
2. LLM interprets and translates the request into a structured query or experiment plan.
3. Search the Gaia-X catalogue for existing data.
4. If found, execute ELT pipeline and deliver dataset.
5. If not found, trigger testbed experiments, collect data, run ELT, and reintegrate into the Data Space.

2.2.2.1.7 Alternative Flow

If the dataset is partially available, the system may recommend augmenting it with additional experiments or transformations before use.

2.2.2.1.8 Post-conditions

The user obtains a high-quality dataset either from existing repositories or from newly executed experiments. The dataset is stored in the 6G Data Space for future reuse.

2.2.2.1.9 High Level Illustration

This illustration refers to the Telefonica Cloud and testbeds use cases. The Telefonica Cloud stands as a key player in digital and data sovereignty, offering an extensive array of services tailored to enhance infrastructure and operational capabilities of businesses.

The high level illustration of the ELT pipeline is shown in the Figure below:
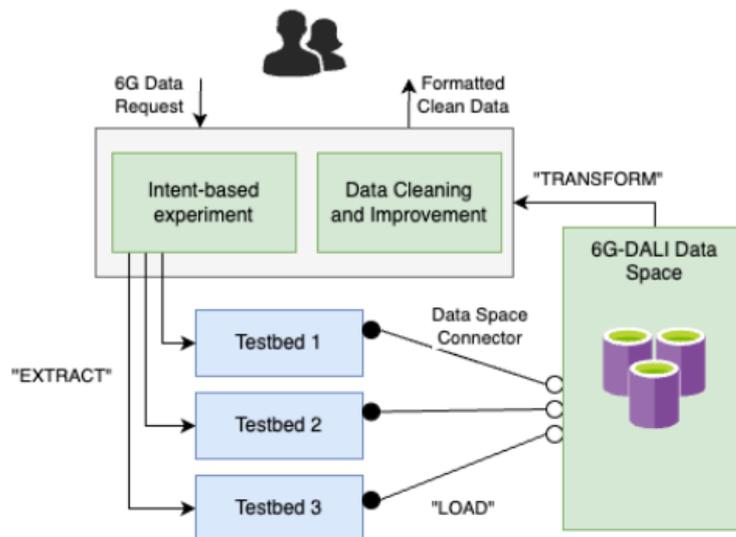
Figure 12: High level illustration of the ELT pipeline

2.2.2.1.10 Potential Requirements

The PoC participants' needs/requirements are the following:

- **Researchers (R) & AI experts and ML developers (AEMD)**, have high needs for robust tools for AI model development, well-labelled and easily discoverable datasets, easy-to-reach execution environments and facilities.

- **Cloud Vendors (CV)**, need scalable, interoperable hosting solutions for ELT workflows and demand seamless integration with data sources.

- **Testbed Business Owner (TBO)**, need testing, orchestration, and monitoring of their facilities

The KPIs of this UC/PoC are the following:

- Experiment: Data on-demand and enhanced ELT pipelines

    o Reduce time to fetch for existing data set by half by report to case where no 6G Data Space exists

    o 100% of data requests expressed via the intent API, based on human language.

    o Showcase 100% compliance with Gaia-X

    o Generate 3 datasets

- Experiment: Data on-demand and enhanced ELT pipelines

- o Improve quality of collected dataset using data cleaning algorithms by 30%

- o Generate 1 new open dataset and perform data cleaning and improvement on 3 existing ones

- o Translate 100% of data requests to experiments, for data generation (i.e., no existing datasets)

- o Satisfy at least 70% of data provisioning requests via the natural language-based APIs.

The requirements for this PoC/Use Case are the following:

REQ 1 Data Space Integration	Enables seamless interoperability with the 6G Data Space, supporting dataset discovery, retrieval, publication, and metadata exchange using standardized interfaces.
REQ 2 Natural Language Query Translation	Leverages an LLM to convert user intents, expressed in natural language, into structured search queries compatible with the 6G Data Space service catalogue.
REQ 3 GAIA-X Trust Compliance	Ensures all data operations adhere to GAIA-X trust, identity, and security frameworks, including authentication, authorization, and policy enforcement.
REQ 4 Curated Dataset Availability	Guarantees access to a baseline of high-quality, pre-processed datasets originating from the project's testbeds, ready for reuse or further transformation.
REQ 5 Auto-Triggered ELT Workflows	Automatically initiates ELT (Extract, Load, Transform) pipelines upon dataset retrieval or generation, streamlining data preparation tasks.
REQ 6 Data Cleaning & Augmentation Support	Integrates mechanisms for cleansing and enriching datasets through domain-specific algorithms, improving dataset quality and model readiness.
REQ 7 Experiment Result Reintegration	Ensures that datasets generated from on-demand experiments are processed and added back into the 6G Data Space for future use by other researchers.
REQ 8 Natural Language to Experiment Translation	Uses an LLM trained to translate high-level user data needs into executable experiment plans, enabling automated testbed utilization.
REQ 9 Multi-Testbed Orchestration	Manages the selection and coordination of multiple 6G testbeds to run distributed experiments efficiently and in parallel, based on user requirements.
REQ 10 Testbed Access Control	Provides secure, programmatic access to the 6G testbeds, enabling controlled experiment deployment, resource allocation, and data collection.
REQ 11 Experiment Monitoring & Logging	Tracks real-time experiment execution across testbeds and logs metadata, status, and output for auditability and debugging.
REQ 12 Quality Assessment Metrics	Defines and applies performance and quality indicators to assess the value of datasets, including improvements from cleaning and augmentation.

2.2.2.1.11 Radio Specific requirements

- Radio Coverage: Multi-cell, indoor/outdoor handover support required

- Bandwidth: High for dataset transfer, packet-based traffic

- URLLC: Latency not critical for ELT but important in live testbed streaming

- Radio Regimens: Licensed spectrum (e.g., testbed specific bands)

- Other: Device location not mandatory; moderate power profile expected

2.2.2.2 Use Case 2: AIaaS for CDN Apps via Cross-Testbed Decentralized MLOps

This use case capitalizes on a meta-orchestration layer that abstracts heterogeneity in MLOps stacks. Each testbed executes ML workflows (training, deployment, drift monitoring) through adapter interfaces. The Hyperparameter Optimization (HPO) engine incorporates AutoML methods such as Bayesian Optimization, while model compression applies quantization-aware training and knowledge distillation to fit edge deployment constraints. The CDN application leverages fs|CDN for benchmarking inference tasks in staging environments, with drift detection enabled via statistical and conformal prediction metrics.

2.2.2.2.1 Description

The second use case (PoC) focuses on the decentralized MLOps capabilities of the 6G-DALI framework, because of the increasing need for robust, automated solutions that can manage machine learning workflows across diverse infrastructures. This PoC aims at validating the functionalities of the 6G-DALI e2e AI framework with decentralized MLOps, meta-orchestration and AI experimentation services, specifically targeting the assessment of its cooperative and distributed capabilities for the management of ML models and showcasing its applicability at a CDN vertical application. In particular, the PoC will validate the 6G-DALI unified approach for managing and executing ML tasks and processes with full automation, streamlining the various steps, and taking autonomous decisions related to AI/ML models training, deployment, placement, hyperparameter optimization, and transfer learning/quantization as a service. Moreover, the PoC aims at assessing the capability of the 6G- DALI e2e AI framework to rely on heterogeneous testbed infrastructures with different MLOps software stacks, demonstrating a solution that is independent from the underlying computing technologies (meta-Orchestration concept) and from the user and ML model constraints.

The objectives of this PoC can be summarized as follows:

- Demonstrate full automation in managing and performing ML processes and workflows, specifically validating streamlined training, deployment, inference, testing and validation tasks.

- Demonstrate the capability to execute AI/ML tasks and processes in a transparent and unified way on top of different testbed infrastructure and MLOps stacks.

- Demonstrate the capability to deploy and place AI/ML tasks and workloads across the extreme- edge, edge, cloud continuum, according to user, data and AI/ML requirements via model compression and quantization.

- Demonstrate the capability to support distributed AI/ML tasks and workloads, such as those required in case of federated learning, cooperative AI/ML inference, and transfer learning.

- Demonstrate the capability to automatically detect model drifts and perform corrective actions according to user, data, and AI/ML requirements.

The 6G-DALI cross-testbed decentralized MLOps PoC is designed to validate the end-to-end AI framework with a focus on decentralized MLOps, meta-orchestration, and AI experimentation services. The PoC targets the evaluation of cooperative and distributed mechanisms for managing machine learning (ML) models and aims to demonstrate their effectiveness within a vertical Content Delivery Network (CDN) use case. It assesses the framework's ability to automate the entire ML lifecycle, covering training, deployment, inference, hyperparameter optimization and transfer learning, without manual intervention. A key goal is to demonstrate that 6G-DALI can operate seamlessly across different testbeds and MLOps stacks, while being able to support distributed AI/ML tasks and workloads.

The motivation behind this PoC is that despite the growing interest in MLOps for automating ML workflows, current solutions often require manual steps at various stages, especially when deployed in complex or production environments. These limitations hinder full automation and scalability, which are essential for realizing the potential of AI in the dynamic, heterogeneous infrastructure characteristics of beyond 5G and emerging 6G networks. Existing MLOps frameworks tend to be fragmented to specific use cases, lacking the flexibility to support diverse

ML techniques and deployment scenarios across the cloud-to-edge continuum. The 6G-DALI PoC addresses these gaps by aiming to offer a unified, adaptable solution capable of overcoming interoperability challenges and enabling consistent AI operations across distributed systems.

The PoC consists of two experiments:

- **Hyperparameter optimization and placement of ML models at the Cloud-Edge-Continuum:** The first experiment will showcase how the experimenters and users can leverage the 6G-DALI framework for efficient ML model training, optimization, and deployment across the Cloud-Edge-Continuum. More specifically, the users can run complex ML workflows, including training, fine-tuning, transfer learning, and model distillation, using the integrated Hyperparameter Optimization (HPO) service, while the system intelligently allocates the minimal computing resources required to meet performance goals like low latency or cost-efficiency. This is particularly valuable for deploying AI workloads on constrained edge or deep-edge devices, where hardware limitations are a key challenge. Users benefit from streamlined experimentation that automatically adapts to infrastructure constraints and optimizes resource use without the need for manual tuning. Additionally, the experiment showcases the use of model compression and quantization, allowing users to deploy powerful AI models even on limited-capacity devices. Overall, the experiment aims to double the number of supported ML workloads using AutoML-driven HPO, ensure deployment of at least one ML workload at a constrained edge device, and generate one dataset as part of the evaluation.

- **ML model Benchmarking and drift detection during vertical application testing:** In the second experiment, the PoC empowers experimenters and verticals to validate, test, and benchmark AI/ML models in staging environments before full production deployment. This capability is demonstrated through a real-world application using ICOM's commercial future-proof CDN (fs|CDN) platform. More specifically, using the 6G-DALI MLOps framework, users can evaluate the performance of AI models that predict resource demand and optimize content delivery, enhancing the Quality of Experience (QoE) for end-users. Experimenters can deploy these models across various cloud-native and edge infrastructures (regardless of ownership or capacity) and use pre-trained models from the ML Catalogue to dynamically control video playback and transcoding settings. Furthermore, the experiment integrates trustworthy AI capabilities, enabling observability and continuous monitoring of AI model behaviour. Users benefit from automated drift detection, which can trigger model retraining or adaptation to maintain accuracy and reliability over time. This ensures that services stay responsive and resilient in dynamic network conditions. The inclusion of data augmentation techniques also helps improve model robustness, allowing users to prepare AI applications for diverse and unforeseen scenarios. Overall, this scenario provides a practical pathway for bringing AI-enhanced services into production, supported by continuous performance monitoring and adaptability features.

2.2.2.2.2 Source

- 6G-DALI Project on CORDIS: https://cordis.europa.eu/project/id/101192750
- 6G-DALI Project portal: https://6gdali.eu/
- 6G-DALI on SNS-JU portal: https://smart-networks.europa.eu/call-3-stream-b/#6G-DALI

2.2.2.2.3 Roles and Actors

This use case (PoC) is expected to involve the following main actors:

- **Vertical Service Provider (VSP)**, representing the stakeholder that delivers domain-specific cases. In the second experiment of this PoC, VSP is ICOM by providing the CDN platform.

- **AI experts and ML developers (AEMD)**, representing the stakeholders that create and optimize the machine learning models through hyperparameter optimization and transfer learning. In the first experiment of this PoC, IBM will deliver the HPO solution representing expertise in developing AI tools, while NXW will deliver the MLOps solution. KUL will contribute to AI model training, particularly in physical layer (PHY) models and data augmentation techniques.

- **Testbed Business Owners (TBO)**, representing the stakeholders that provide the testbed infrastructure where the experiments are deployed. KUL and EUR will provide the testbeds that will be involved in the experiments.

- **Cloud Vendor (CV)**, representing the stakeholder that hosts the 6G-DALI framework including ML models catalogue and services for training and inference. The 6G-DALI framework will be hosted by ATH/ISI infrastructure/cloud.

The participants benefit from the PoC in the following ways:

- **Vertical Service Providers (VSP)** will gain predictive capabilities for content/resource management and will benefit from MLOps-driven automation and QoS optimization in their services.

- **AI experts and ML developers (AEMD)** will leverage the AutoML-driven HPO services to successfully deploy models at constrained edge devices.

- **Testbed Business Owners (TBO)** validate their infrastructure's capability with heterogeneous MLOps stacks.

- **Cloud Vendor (CV)** demonstrate value as infrastructure providers for scalable, distributed AI workloads across cloud-native and edge environments.

2.2.2.2.4 Pre-conditions

- ML models exist and need optimization or testing.

- MLOps environments are operational across distributed infrastructures.

- CDN use case scenario is defined.

- Benchmarking of pre-trained models is available.

- Model compression is enabled.

2.2.2.2.5 Triggers

Requests from verticals to train, test or benchmark AI models, deploy models across cloud-to-edge within resource constrained devices, or detect performance drift and trigger re-training if needed.

2.2.2.2.6 Normal Flow

1. User initiates training or benchmarking via API.

2. Meta-orchestrator determines optimal deployment across testbeds.

3. HPO service optimizes models.

4. Models are placed using compression/quantization.

5. Drift monitoring ensures continued accuracy, with retraining if needed.

2.2.2.2.7 Alternative Flow

If target infrastructure cannot support the full model, automated compression or transfer learning is triggered to adapt the model for constrained devices.

2.2.2.2.8 Post-conditions

- Optimized and deployed AI models enhance CDN services.

- Performance improvements are measured and stored.

- Updated models and datasets are saved in the ML catalogue.

- Compressed models are suitable for deployment on a specific target device.

- Updated models after drift detection meet predefined baseline performance or improvement threshold.

2.2.2.2.9 High Level Illustration

This use case will use testbed facilities from KUL and ATH/ISI.

The KUL testbed can be sued for Real-time Flexible Radio Access Network research, offering to experimenters:

- Remote data logging capability to sense channel data for long periods of time.

- Remote (ssh) access of the data server for training AI models with an Nvidia A100 RTX GPU (more streamlined MLOps instrumentation deployed within SUNRISE-6G).

- Remote access to the radios to trigger remote experiments (more streamlined testing-as-a-service workflow implemented within 6G-BRICKS).

The infrastructure of our testbed consists of a main computer, to perform computation and control, and a set of Universal Software Radio Peripherals (USRPs), which can operate as Radio Units (RUs). This main computer is an HP Enterprise Edge server configuration featuring a dual Xeon Gold 6430 processors boasting 32 cores and 64 threads in total. The machine has 1 TB of DDR5 PC5-4800-2R memory and 960GB SATA SSD storage. The connectivity between the controller and the RUs is provided by a dedicated high-speed 100GbE fibre link. Furthermore, a 1Gb Ethernet interface is designated for control purposes.
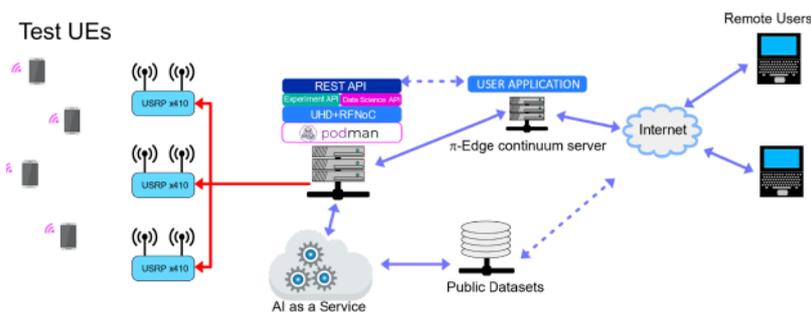
Figure 13: Higher illustration of AIaaS for CDN Apps via Cross-Testbed Decentralized MLOps

The main computer not only provides the core operational aspects of the testbed but also the software and hardware for machine learning, and data processing. Users can perform number

crunching and AI training using an NVIDIA A100 accelerator card, in conjunction with Tensor Flow or PyTorch. Data sets are stored locally and are accessible to the public, ensuring transparency and availability. To ensure precision in timing and synchronization, a Rubidium clock governs the synchronization of the entire testbed. The testbed is also able to distribute the clock and synchronize the different components using IEEE1588/PTP.

The radio units are USRP X410 by National Instruments. They are capable of operating either in a stand-alone (embedded) mode or as host-based systems with network streaming capabilities. These units cover a frequency range from 1 MHz to 7.2 GHz, with tunability extending up to 8 GHz. The RUs have an instantaneous bandwidth capacity of up to 400 MHz.

Accurate positioning of the user is also ensured by using a motion capture system (Qualisys Miqus M3 with 4 cameras) for sub-mm position. Cameras are arranged all around the room but can be moved around to capture different scenarios. The cameras have 2Mpixel (1824 × 1088 resolution) and capture IR video at 340 fps. They can correctly resolve positions with 0.11mm precision in 3D. The maximum range is 15m. The system comes with a dedicated software to capture data, but it is compatible with MATLAB and Python. Data capture can be synchronized to the same time reference used by the testbed.
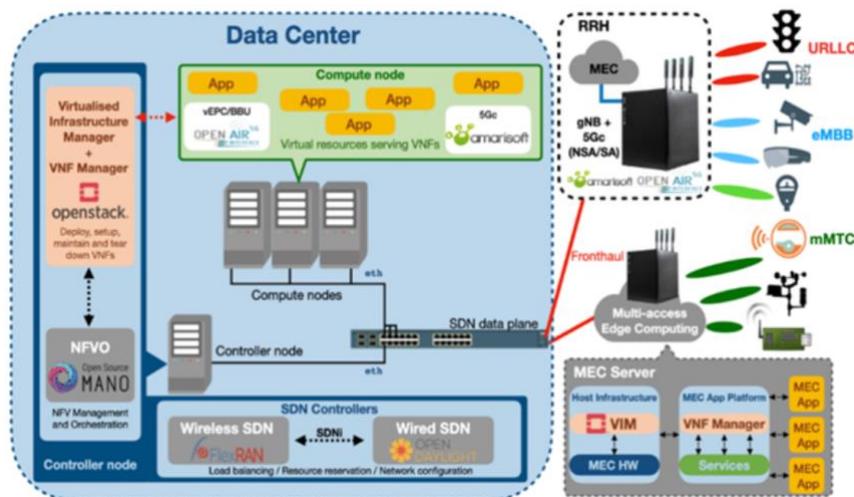
Figure 14: Illustration of the ATH 5G testbed facility

The ATH 5G testbed facility, located at the ISI premises in Greece, implements a full stack 5G system, encompassing Core, Edge and RAN tiers as well as advanced 5G UE emulation capabilities. The ATH testbed is based on open-source frameworks and software stacks, including an Amarisoft 5G RAN and Core Network, Kubernetes/ OpenStack based MEC nodes and ETSI OSM based Orchestration components. It is currently being upgraded with O-RAN compliance while zero-touch management, MANO and dynamic slice reconfiguration enablers are developed.

The ATH testbed offers its resources via a Cognitive Compute Continuum enabler. AI optimization algorithms are leveraged throughout the management of computing, networking and storage resources located anywhere in the network and therefore defining an expanded network-compute fabric facilitating a hyper-distributed and hyperconnected environment. This is facilitated by an abstraction framework which is implemented within 6G-BRICKS to facilitate the dynamic deployment, instantiation, migration, and scaling of cloud-native service chains, following their constantly evolving networking and compute requirements and the changes in the underlying infrastructure (e.g., new resources becoming available).

Dataset capabilities: ATH testbed supports the generation of datasets from the Cognitive Cloud-Edge deployment via k8s telemetry APIs (e.g., related to CPU utilization, RAM, pod placement, CNI). Moreover, dataset generation through O-RAN telemetry will be available through the O-RAN E2 interface and Service Models (SM).

MLOps and RLOps support to 6G-DALI: The ATH testbed is part of SUNRISE-6G and ADROIT-6G, where MLOPS tools have been adopted for automated AI model lifecycle management via MLFow.

2.2.2.2.10 Potential Requirements

The use case (PoC) participants' needs are the following:

- **Vertical Service Providers (VSP)** need to use 6G-DALI as a staging/testing framework that mimic production CDN use case to compare various models' performances and ensure continuous monitoring of the deployed model's performance.

- **AI experts and ML developers (AEMD)** need advanced tools for model lifecycle management (training, tuning, inference) to support their activities.

- **Testbed Business Owners (TBO)** need testing, orchestration, and monitoring of their facilities.

- **Cloud Vendor (CV)** need scalable, interoperable hosting solutions for diverse AI workflows and demand seamless integration with diverse MLOps stacks.

The KPIs of this use case (PoC) are the following:

- **Experiment: Hyperparameter optimization and placement pf ML models at the Cloud-Edge-Continuum**

  o Double the number of ML workloads supported via HPO driven by AutoML.

  o Deploy at least 1 ML workload at a constrained edge device.

  o Generate 1 dataset.

- **Experiment: ML model Benchmarking and drift detection during vertical application testing**

  o Have at least 0.5 MOS score gain for the CDN via resource prediction.

  o Reduce at least by 20% vertical SLA breaches via trustworthy AI techniques such as drift detection and conformal prediction.

  o Improve by at least 20% prediction accuracy for the PHY model in unseen scenarios, via Trustworthy AI approaches that improve model generalization ability via data augmentation.

  o Generate 2 datasets.

The requirements of this use case (PoC) are the following:

REQ 1 Model Benchmarking	Enables benchmarking of pre-trained models available in the ML catalogue.
REQ 2 Model Compression	Enables model compression to fit them within resource constrained devices. Includes configurable parameters such as the compression technique (pruning / quantization / knowledge distillation / a combination from the catalogue), the compression level (manually defined / based on target accuracy / auto-optimized), the validation strategy (available in the catalogue / custom-defined / automatically selected based on the model and task), and the evaluation metric (available in the catalogue / custom-defined / automatically selected).
REQ 3 Model Drift Detection / Continuous Monitoring	Enables continuous monitoring of performance of deployed ML models to detect any degradation. Includes configurable parameters for the drift detection process such as the drift metric (statistical distance measure / distribution shift metric / learned or heuristic-based metric) and the detection frequency (real-time / batch / daily). In case of deviation, either from a performance threshold or from a baseline model, model re-training is triggered to enhance robustness.

2.2.2.2.11 Radio Specific requirements

- Radio Coverage: Local/multi-cell coverage depending on deployment site

- Bandwidth: Medium to high depending on model/data size

- URLLC: Required for inference at edge; low latency target

- Radio Regimens: Licensed/private networks for testbed realism

- Other: Power-efficient models for edge; location services not critical

2.2.2.3 Use Case 3: DTT and RLOps for Large and Medium-Scale Experiments

The Digital Twin Testbed (DTT) is built atop Open-Air Interface (OAI) and VIAVI test systems. Its configuration layer defines topologies, RAN/CN parameters, and UE profiles. Experiments are defined using templates and executed using xApp logic or NWDAF agents. The system supports low-fidelity simulation for scale (thousands of UEs) and high-fidelity radio interaction for O-RAN testing. Reinforcement Learning agents are onboarded into the RIC and evaluated using KPIs such as handover efficiency and CN function scalability. Datasets are ingested into the 6G Data Space post-ELT.

2.2.2.3.1 Description

This use case leverages the Digital Twin Testbed (DTT) and Reinforcement Learning Operations (RLOps) to enable medium- and large-scale experiments for 6G research. It addresses the scarcity of realistic, large-volume datasets for training AI and RL models in future mobile networks. The DTT integrates with the 6G Data Space using ELT pipelines and supports automated experiment configuration, data collection, and agent validation. This PoC showcases 6G-DALI capabilities to run large-scale experiments using the AI framework and the DT, with two-fold objectives. First, datasets are generated and integrated into the 6G dataspace, using 6G-DALI ETL. Second, realize RLOps to test and validate trained RL agents. The objectives of this PoC are:

- Demonstrate the automatic deployment of large-scale experimentation on the 6G-DALI from the definition, deployment, and generation of datasets involving the AI framework components: user gateway, data acquisition, data cleaning, adaptation layer, and DTT adapter.

- Demonstrate the configuration and data set collection using the OEM DTT and xApp.

- Demonstrate the DT capabilities to generate Medium and large-scale experimentations with low and high-fidelity DT.

- Demonstrate the integration of the DTT via ELT to the 6G Data Space.

- Demonstrate RLOps mechanisms devised in 6G-DALI.

The growing need to develop machine learning models tailored for 6G networks is significantly constrained by the lack of representative datasets, particularly when it comes to modelling the RAN and CN. This limitation arises not only from the scarcity of publicly available datasets but also from the inadequacy of the scenarios used to generate them. Most existing datasets fail to capture the scale and complexity expected in 6G environments, such as a high density of UE per cell, large-scale multi-cell deployments, and realistic mobility and handover scenarios. While many existing platforms like those introduced by SNS-C projects (6G-Bricks[1], 6G-Sandbox[2], and Sunrise-6G[3]) facilitate experimentation with different configurations for 6G RAN and CN, they are typically limited in the number of UEs that can be simulated or emulated. This restricts the ability to evaluate critical features such as Xn and F1 handovers, as well as to assess the scalability of CN functions and algorithms designed to manage dense UE populations. Furthermore, many of these platforms do not support the execution of machine learning models or reinforcement learning agents directly on the testbed. This hinders the validation of such models in realistic, scalable environments—an essential requirement for ensuring their practical applicability in 6G networks. Moreover, many platforms do not support the execution of ML models and RL agents, which require validation in realistic and scalable environments that reflect genuine deployment scenarios. Knowing these limitations, 6G-DALI devises a novel Digital Twin Testbed (DTT) for the 6G RAN and CN. The DTT will allow the running of experiments and, hence, dataset generation for scenarios that require a high-scale deployment running a high number of mobile UEs, several gNB cells, different RAN configurations, etc.

This DTT testbed is essential for evaluating next-generation network technologies before they are deployed in the real world. The DTT enables large-scale experimentation by simulating thousands of mobile devices and network elements such as base stations and core networks. This allows researchers and engineers to test how different network configurations perform under various conditions, including high traffic loads and dynamic user movements. The insights gained from these experiments help shape more efficient, reliable, and adaptable 6G networks. The testbed utilizes both open-source and commercial tools. It leverages the well-established Open-Air Interface [4] (OAI) 5G tools in EURECOM and testing tools from VIAVI that validate how well different network components work together. These tools can simulate real-world usage, check for compliance with industry standards, and ensure that the network performs well under pressure. Additionally, the system is designed to monitor and adjust network operations intelligently. It uses specialized applications that gather performance data and make real-time decisions to optimize how the network functions. This dynamic control is a key element for future smart networks that can adapt to changing conditions instantly.

The DTT setup includes three main components:

1. Configuration manager defines experiments, e.g. network layout, device behaviour, traffic patterns.

2. Intelligent control layer that monitors network performance and adjusts as needed.

3. Adapter connects all components to a broad AI framework, automating experiments & data collection.

All experimental data is stored and shared via a centralized 6G data platform, allowing for further analysis and training of AI models. This collaborative and automated approach accelerates the development of innovative 6G technologies and ensures they are tested rigorously before real-world implementation.

In this use case where 6G-DALI will be executed on DTT, we envision two scenarios:

1.   Large-scale experimentation for RLOps

2.   Medium-scale DT experimentation for O_RAN

Both scenarios have common objectives, which demonstrate:

- Automatic deployment of large-scale experimentation on the 6G-DALI from the definition, deployment, and generation of datasets involving the AI framework components: user gateway, data acquisition, data cleaning, adaptation layer, and DTT adapter.

- Configuration and data set collection using OEM DTT and xApp, and application run on top of NWDAF.

- DT capabilities to generate Medium and large-scale experimentations with low and high-fidelity DT.

- The integration of the DTT via ELT to the 6G Data Space.

- RLOps mechanisms devised in 6G-DALI.

Large-scale experimentation for RLOps

Large scale experimentation for RLOps, noted also as large-scale and low fidelity, aims to experiment with scenarios that involve deploying a high number of UEs connected to different cells while supporting different mobility patterns. In this scenario, the UEs and RAN are based on VIAVI tools, while the RIC and CN use OAI. This scenario has a lower fidelity as no radio is used. However, this scenario support large scale experiments, by enabling the collection of data on user mobility, high number of attach requests, and high traffic patterns, which can be used to train RL agents that optimize the system performances, such as mobility and handover procedures, dimensioning of CN functions, etc. The RL agents are expected to be on-boarded as xApp on top of RIC or an application running on top of the NWDAF. This will pave the way to showcase 6G-DALI RLOps system in terms of its SIM2REAL generalization capabilities.

This scenario has these specifics objectives:

- Demonstrate the DTT capabilities to generate large-scale experiments.

- Demonstrate RLOps system of 6G-DALI.

- Showcase 6G-DALI ELT capabilities.

Medium-scale DT experimentation for O-RAN

The second scenario focuses on testing with a moderate number of real user devices using actual radio signals. This setup uses equipment from VIAVI to simulate realistic mobile network conditions. Unlike the first scenario, which was fully virtual, this one includes real radio hardware to closely mirror real-world deployments. The goal is to collect valuable data on how the network performs across different parts, including user connections, central network behaviour, and how devices move between network areas. This type of data is not yet available in O-RAN environments, as the technology is still new and current testing platforms are limited.

To support this, custom applications will automatically collect and organize the data during the experiment as part of the 6G-DALI project. This will help demonstrate what is possible with more advanced and realistic O-RAN testing.

The specific objectives of this scenario are:

- Demonstrate DTT capabilities to generate medium scale experiment with high fidelity targeting complex scenario such as O-RAN.

- Showcase ELT capabilities.

- Demonstrate experiment on demand using DTT.

2.2.2.3.2 Source
- 6G-DALI Project on CORDIS: https://cordis.europa.eu/project/id/101192750
- 6G-DALI Project portal: https://6gdali.eu/
- 6G-DALI on SNS-JU portal: https://smart-networks.europa.eu/call-3-stream-b/#6G-DALI

2.2.2.3.3 Roles and Actors
The successful execution of this use case (PoC) within the 6G-DALI framework relies on the collaboration of several key participant roles. Each plays a distinct function to support the development, deployment, and evaluation of both large-scale RLOps experimentation and medium-scale DT experimentation for O-RAN. The roles are as follows:

- **AI Experts and ML Developers (AEMD):** Design, implement, and tune RLagents. They integrate AI framework components to automate experimentation and ensure seamless data processing.

- **Testbed Business Owners (TBO):** Provide and manage the physical and virtual infrastructure for experimentation. This includes setting up the OEM DTT, enabling xApp deployment, and supporting the execution of scenarios via NWDAF for data collection and real-time feedback.

2.2.2.3.4 Pre-conditions
- Digital Twin Testbed components are deployed and configured.

- RL agents are ready for testing and evaluation.

- The AI framework, NWDAF/xApp interfaces, and ELT pipeline are functional.

2.2.2.3.5 Triggers
A user submits an experiment request to evaluate RL models at scale or to generate data for O-RAN compliant scenarios.

2.2.2.3.6 Normal Flow
1. Define experiment parameters and scenarios in DTT configuration manager.

2. Automatically deploy and execute experiments using simulated or real radio environments.

3. Use ELT pipeline to clean, augment, and store datasets.

4. Evaluate RL agents' performance through xApp or NWDAF.

5. Share results and datasets via the 6G Data Space.

2.2.2.3.7 Alternative Flow

If high-fidelity radio hardware is not available, fallback to large-scale, low-fidelity simulation mode for RLOps experimentation.

2.2.2.3.8 Post-conditions

Validated RL models and multiple large-scale datasets are produced and available in the 6G Data Space.

2.2.2.3.9 High Level Illustration

The EURECOM testbed shall be used in this use case, presented in the figure below.
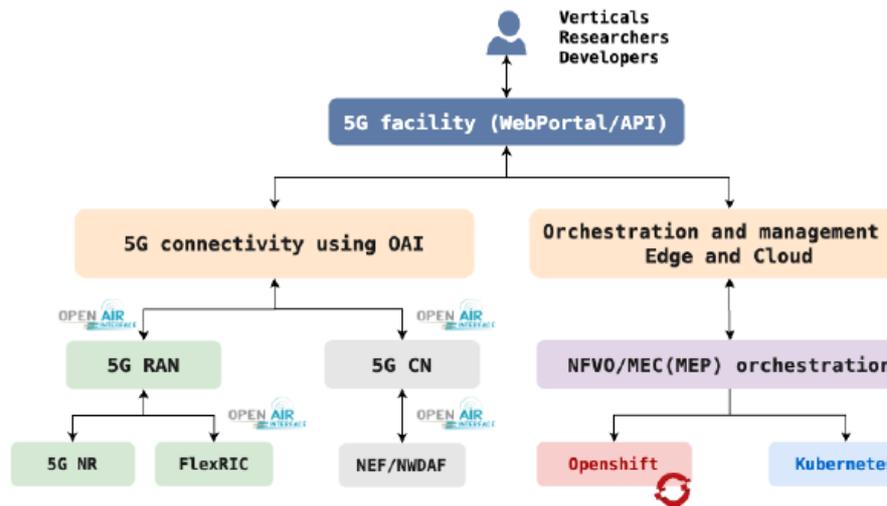
Figure 15: The EURECOM testbed

The 5G facility of EUR, as illustrated in the above Figure, provides experimental 5G services, including eMBB, URLLC, and mMTC. Based on fully open-source tools and open-architecture design, it provides the means to on-board new network functions to the running 5G infrastructure and test them in both a controlled laboratory setting and in a deployed live network. As shown in the figure above, the facility accessible through a Web Portal or exposed API. The Web Portal is specifically designed to allow vertical use-case deployment, while the API can be use by an external component of orchestrator. Indeed, the facility is connected to EU federation of testbed in the context of the SLICE-PP, 6G-BRICKS and SUNRISE-6G projects.

In terms of technical components, the facility provides 5G connectivity using Open Air Interface (OAI). It permits the deployment of cloud-native applications and services in the forms of Virtual Network Function (VNF) or Container Network Function (CNF). The management and orchestration components are compliant with the ETSI NFV and MEC specifications to manage virtual resources for cloud and edge. The NFVO and MEC Orchestrators are homemade software that allows the manage the Lifecycle of Network Services via Network Service Descriptor (NSD), which in turn is composed of a set of cloud-native functions (i.e., CNF) described using an Application Descriptor (AppD).

The cloud-native functions are deployed on two container-management platforms, i.e., OpenShift and Kubernetes. The 5G NR connectivity is fully supported by Open Air Interface (OAI) RAN and CN. The RAN supports both FR1 and FR2 deployments. Regarding FR1 deployment, the RAN platform is split into CU, DU, and RU, following two different functional splits.

- **Split 7.2**: OAI CU & DU use commercial RUs from VVDN (indoor), LiteON (indoor), Benetel (indoor/outdoor)

- **Split 8**: OAI CU & DU using USRPs (B210, N300, X310, indoor) and commercial AW2S RUs (indoor/outdoor)

The FR1 RAN setup operates at bands 38 (2.6 GHz) and 78 (3.4 GHz) supporting two subcarrier-spacing (15 and 30kHz), six different bandwidth sizes (10, 20, 40, 50, 80 and 100MHz) with multiple BWPs support (initial BWP and dedicated BWPs), different TDD configurations enabling asymmetric assignment of uplink and downlink resources with shorter TDD periods (down to 2.5 ms). It also includes the Procedures for 4-layer DL and 2-layer UL MIMO with the support of 256 QAM modulation. For FR2, the current implementation supports FR2 RU from a commercial company, LiteON, operating at 27 GHz. Completing the O-RAN setup, the facility includes a Real-time Intelligent Controller (RIC) (OAI FlexRIC) to run xApp for RAN management and optimization.

Regarding the 5G CN, it is based on OAI CN. It is a partial 3GPP 5GC service-based architecture including the following Network Functions (NFs): NRF, AMF, SMF, UPF, UDM, UDR, AUSF, NSSF, PCF, NEF, NWDAF and CAPIF. NEF and CAPIF are used to expose CN functions to verticals.

2.2.2.3.10 Potential Requirements

AEMDs require robust access to AI tooling, clean and well-labelled datasets, and reproducible experimentation environments that facilitate continuous development.

TBOs need flexible infrastructure capable of supporting low- and high-fidelity Digital Twin deployments, along with tools for real-time monitoring and orchestration.

Participants' benefits include:

AEMDs benefit from an automated experimentation pipeline, faster model validation cycles, and access to rich, high-quality datasets for training and evaluation.

TBOs strengthen the value of their testbed infrastructure by supporting cutting-edge use cases, increasing visibility, and fostering collaboration across the 6G ecosystem.

The KPIs of this POC are the following:

- Large-scale experimentation for RLOps:

    o Reduce by 50% the time required to validate a reinforcement learning agent compared to scenarios without Digital Twin Technology.

    o Ensure that 80% of experiments are executed automatically (no human in the loop).

    o Generate at least two large datasets tailored for large-scale scenarios.

- Medium-scale Digital Twin experimentation for O-RAN:

    o Ensure that 80% of experiments are automated (no human in the loop).

    o Generate at least two large datasets for medium-scale scenarios, fully compliant with O-RAN specifications.

    o Ensure that 100% of the generated data is shared with various stakeholders via the 6G Data Space.

2.2.2.3.11 Radio Specific requirements

- Radio Coverage: Medium (O-RAN), Large (DTT simulation); multi-cell needed

- Bandwidth: High for telemetry and dataset generation

- URLLC: Low latency required for real-time agent validation

- Radio Regimens: Combination of licensed and emulated spectrum use

2.2.2.3.12 Other requirements

- Requires continuous monitoring; battery

2.2.3 6G-INTENSE (Intent-driven NaTive AI architecturE supporting Compute-Network abstraction and Sensing at the Deep Edge) Use Cases

6G-INTENSE (https://6g-intense.eu/) proposes a new System Architecture for 6G, to deliver "6G as a Smart Service Execution platform", fully in line with the vision of sustainable infrastructure sharing to reduce space and energy costs, and encouraging collaboration among all members of the value chain under a unified Network-Compute fabric. Key contribution is a novel automation architecture with a Native AI toolkit facilitating intent declaration, negotiation and decision automation across autonomous domains, termed as Distributed Intent-driven Management and Orchestration (DIMO). The new business model introduced by 6G-INTENSE separates service management from resource management and creates a unique opportunity to disrupt the market to allow the entrance of new actors, such as those owning extreme-edge nodes.

2.2.3.1 Use Case 1: Distributed Continuum towards Pervasive Computing

The first Use Case concerns the Distributed Continuum towards pervasive computing. It aims at validating the Abstraction Framework that will be developed in the context of 6G-INTENSE and will address the deployment, model training and integration with the southbound Compute and Communication platforms provided by different partners, namely the **πEdge** cloud-native edge management platform from Intracom Telecom, **OpenAirInterface Open RAN** from EURECOM and **MPLS** from Orange Romania. The Network-Compute Fabric introduced by 6G-INTENSE aims to provide an advantage to telecom operators in terms of increased efficiency through joint optimization and allocation of network and computational resources. At the same time, network infrastructures can act as hosting environments for application tasks and (micro-) services, allowing ubiquitous computing and extreme proximity of user-centric processes to end user environments as well as localized inference and learning, calibrated by converged sensing features of the (Deep) Edge.

Content consumption and gaming represent the most critical use-cases that drive internet traffic and are rapidly shifting towards mobile devices. Mobile applications have a strong dependence on user location and typically depend on Content Delivery Networks (CDNs), which are being used to reduce the delivery latency of content requested by subscribers. A CDN typically has infrastructure in multiple geographical locations, known as Points of Presence (PoPs). Each PoP contains several storage, caching, and streaming servers responsible for content delivery of the live and on-demand content to users within its proximity. Therefore, a subscriber of a CDN service can receive content from the corresponding serving PoP, depending on their location. In case of fixed clients (set-top boxes, smart TVs, etc.), there is a one-to-one mapping of a subscriber to a PoP. However, with emergence of smartphones and tablets as end devices, a subscriber can connect to various PoPs as they move across respective mobile network interfaces.

Use Case 1 will support two different experiments validating the 6G Network-Compute Fabric abstraction framework that will be developed in the project, leveraging on the Media Vertical to showcase the deployment and operation of Content Delivery Networks in the form of micro-services on a variety of heterogeneous resources and infrastructures, unified under the umbrella of Cloud Edge Continuum.

2.2.3.1.1 Use Case 1 - Scenario 1: Pervasive Computing in a distributed continuum

2.2.3.1.1.1 Description

This experiment concerns scenarios about the smart and efficient distribution of media content (video streaming, gaming, AR/XR, etc.) across a multi-domain, multi-cluster Core/Cloud to Edge to Deep Edge Continuum. To this end, we consider a resource layer of converged networking and compute infrastructure via the Network-Compute fabric capable of hosting a hierarchical,

dynamic set of Edge Cache instances to support the reliable and efficient streaming of content over a geographically distributed customer base.

During the execution of the experiment, we will showcase the instrumental role of the 6G-INTENSE Network-Compute Abstraction framework, homogenizing in an abstracted view all underlying resources and allowing for the hierarchical, intent-driven automation. Thank to that the optimal Edge Cache services' instantiation, configuration, operation, and adaptation (auto-scaling, auto-healing, migration etc.) will be able to respond to the sensed context of evolving environments and users. Moreover, this experiment targets the lifecycle management of micro-service deployment as an Orchestration Continuum undertaken by the Network-Compute Fabric abstraction framework after appropriate training of its Generative AI models. The JCS of the Deep Edge infrastructure (e.g. towards the monitoring of the network links and performance) will be another outcome of the experiment.

2.2.3.1.1.2 Source

- 6G-INTENSE at CORDIS: https://cordis.europa.eu/project/id/101139266
- 6G-INTENSE project page: https://6g-intense.eu
- 6G-INTENSE SNS JU web: https://smart-networks.europa.eu/phase-2-stream-b/#6G-INTENSE

2.2.3.1.1.3 Roles and Actors

The primary roles/actors in this use case include:

- Verticals: Entities that aim to deploy their applications or services as a 6G service on top of the 6G infrastructure.

- Media Providers: Offering content-heavy services that benefit from edge delivery.

- Telecom Operators/DMOs: Managing the lifecycle of 6G services deployed on the 6G infrastructure.

- Infrastructure / resource providers: Own one or more different types of 6G resources.

- End users: Consuming streaming and gaming services, whose location dynamically affects PoP association.

Relationships between them centre around the dynamic provisioning, orchestration, and optimization of services based on user demand and mobility.

2.2.3.1.1.4 Pre-conditions

The distributed infrastructure must be available and ready for use, including the southbound platforms: πEdge from Intracom Telecom, Open RAN from EURECOM, and IP/MPLS networks from Orange Romania. The fs|cdn® Anywhere platform must also be registered with the orchestrator, and AI models trained to optimize service placement must be available. Connectivity among edge nodes and the core should be provisioned to support intent propagation and service deployment.

2.2.3.1.1.5 Triggers

This use case is triggered when a tenant initiates a service request for content delivery, such as a video stream. The request is expressed in natural language (e.g., deliver a live stream to a specific region) and is translated to a high-level service intent as input to the DMO. The DMO then translates the service intent to resource intent and provides it as input to the Network-Compute Fabric, which converts the resource intent into specific configurations and actions for service deployment and management.

2.2.3.1.1.6 Normal Flow

1. Vertical requests CDN service instantiation in natural language.

2. Request is converted into service intent and sent to DMO.

3. DMO translates service intent into resource intent and forwards to Network-Compute Fabric.

4. Network-Compute Fabric translates resource intent into specific requests for the involved southbound platforms and optimally deploys the appropriate CDN components.

5. User receives content from optimal PoP.

2.2.3.1.1.7 Alternative Flow

Fallback to centralized CDN if edge deployment is not viable. If sufficient edge resources are not available or network congestion prevents optimal placement, services may be deployed closer to the core, with degraded SLA (e.g., increased latency). AI-based fallback strategies attempt to reallocate resources or reroute traffic dynamically.

2.2.3.1.1.8 Post-conditions

Once the streaming session ends or user mobility shifts the context, the services are either scaled down or migrated. Service lifecycle concludes with cleanup and resource release. Logs and telemetry are retained for KPI evaluation and future model retraining.

2.2.3.1.1.9 High Level Illustration

This scenario will utilize ICOM's fs|cdn™ Anywhere commercial solution, its own-developed and future-proof IPTV & multi-play services platform, shown in **Figure 16: Illustration of fs|cdn™ Anywhere**. ICOM's end-to-end fs|cdn® Anywhere platform enables service providers to deliver high-quality video to any television, computer, smartphone, or tablet by smoothly bundling middleware and conditional access, offering a complete portfolio of value-added service packages (e.g. live TV, video on demand, restart TV, network DVR and TV) anywhere over any managed or unmanaged network. It combines a fully customizable EPG and advanced TV features to build an engaging, locally-branded TV service and makes access to services ubiquitous across any device. fs|cdn® Anywhere products work together seamlessly, so subscribers can move between home TV service, mobile viewing, and on-demand/DVR services with ease.

Figure 16: Illustration of fs|cdn™ Anywhere

2.2.3.1.1.10 Potential Requirements for Use Case 1 - Scenario 1: Pervasive Computing in a distributed continuum

Content consumption and gaming represent the most critical use-cases that drive internet traffic (e.g., Netflix uses up 15% of all Internet downstream traffic worldwide), and are rapidly shifting towards mobile devices. Mobile applications have a strong dependence on user location and typically depend on CDNs with global PoP deployments. This market is currently dominated by US-based hyperscalers and is difficult to disrupt by EU companies due to the global nature of CDNs.

The requirements for this experiment are summarized in **Table 17: Requirements for Use Case 1 Scenario 1 (Pervasive Computing in a distributed continuum)**.

Table 17: Requirements for Use Case 1 Scenario 1 (Pervasive Computing in a distributed continuum)

ID	UC1–S1-01
Name	fs|CDN Service Instantiation
Description	User must be able to request instantiation of the fs|CDN service.
ID	UC1–S1-02
Name	Requirements specification
Description	User must be able to specify service-related constraints and requirements/targets.
ID	UC1–S1-03
Name	Service update
Description	User should be able to update their service request.
ID	UC1–S1-04
Name	Service termination
Description	User should be able to terminate their instantiated service.

2.2.3.1.1.11 Radio Specific Requirements

Radio Coverage:

- <u>Cell range</u>: local (edge site), urban indoor/outdoor

- <u>Multicell support</u> required

- <u>Seamless handover</u> with <5ms interruption tolerance

- <u>Typical user speed</u>: pedestrian (~5 km/h), vehicular (<50 km/h)

2.2.3.1.1.12 Bandwidth:

- <u>Peak rate</u>: >100 Mbps

- <u>Average rate</u>: 30-50 Mbps

- Packet-switched, bursty traffic

2.2.3.1.1.13 URLLC:

- <u>One-way latency</u> <10ms

- <u>Reliability</u> > 99.999%

- <u>Max jitter</u>: 5ms

2.2.3.1.1.14 Other requirements:

- <u>Energy efficiency</u> critical for edge nodes

- <u>Terminal location accuracy</u>: ~1 meter for optimal cache placement

2.2.3.1.2 Use Case 1 - Scenario 2: Edge Intelligence and Compute Interconnection

2.2.3.1.2.1 Description

6G systems approach for the increasing importance CDNs for high-quality video streaming is based on the distributed and scalable architecture of DIMO and Network Compute Fabric techniques. The aim is to ensure Service Level Agreements (SLAs) for CDN-based video streaming services deployed over different domains, validated through a CDN-based video service experiment, where a high-level defined intent is translated and propagated across the Network-Compute Fabric, connecting computing and RAN domains. The experiment focuses on the Composable AI and learning capabilities of the 6G-INTENSE system, utilizing Machine Learning as a Service (MLaaS) to automate objective setting and adjustment across domains and resources.

The goal is to optimize streaming services by transferring learning assets between domains and conducting domain adaptation. The experiment aims to demonstrate the usage of SD-WAN to facilitate CDN PoPs interconnection as well as the dynamic resource integration to the Network-Compute Fabric abstraction framework. Another goal is the SLA guarantee by employing intent monitoring techniques and the Edge intelligence for training distributed ML models. The scenario objectives are also related to the chaining concepts under a unified entity, the Network-Compute fabric.

More specifically, to facilitate intent declaration, negotiation, and decision automation across Edges domains, as described in **Figure 17**, in relation to the end-user (Use Cases experiments) requirements, dynamic service and resources are orchestrated with the support of the SD-WAN.

2.2.3.1.2.2 Source

- 6G-INTENSE SNS JU web: https://smart-networks.europa.eu/phase-2-stream-b/#6G-INTENSE
- 6G-INTENSE at Europa portal: https://cordis.europa.eu/project/id/101139266
- 6G-INTENSE home WEB page: https://6g-intense.eu

2.2.3.1.2.3 Roles and Actors

The primary roles/actors in this use case include:

- Verticals: Entities that aim to deploy their applications or services as a 6G service on top of the 6G infrastructure.

- Media Providers: Offering content-heavy services that benefit from edge delivery.

- Telecom Operators/DMOs: Managing the lifecycle of 6G services deployed on the 6G infrastructure, enhancing SLA assurance and dynamic PoP configuration.

- Infrastructure / resource providers: Own one or more different types of 6G resources, integrating programmable SD-WAN to link RAN and Edge.

- Edge AI Providers: deploy and train distributed ML models.

- End users: Benefit from uninterrupted video services with improved quality.

2.2.3.1.2.4 Pre-conditions

Distributed CDN services and Network-Compute Fabric abstraction with integrated SD-WAN controllers.

2.2.3.1.2.5 Triggers

Service intent declaration, resource availability, SLA degradation alerts.

2.2.3.1.2.6 Normal Flow

1. Intent for video service occurs from Tenant request.

2. DMO translates service intent into resource intent and forwards to Network-Compute Fabric.

3. Network-Compute Fabric translates resource intent into specific requests for the involved southbound platforms and provisions CDN PoPs.

4. SD-WAN interconnects domains dynamically.

5. AI models adapt streaming quality and resources.

2.2.3.1.2.7 Alternative Flow

Fallback to centralized orchestration if SD-WAN paths are unavailable.

2.2.3.1.2.8 Post-conditions

Video service operates with guaranteed SLA; logs and metrics update AI training datasets.

2.2.3.1.2.9 High Level Illustration

The experiment focuses on the Composable AI and learning capabilities of the 6G-INTENSE system, utilizing Machine Learning as a Service (MLaaS) to automate objective setting and adjustment across domains and resources, as in **Figure 17**.
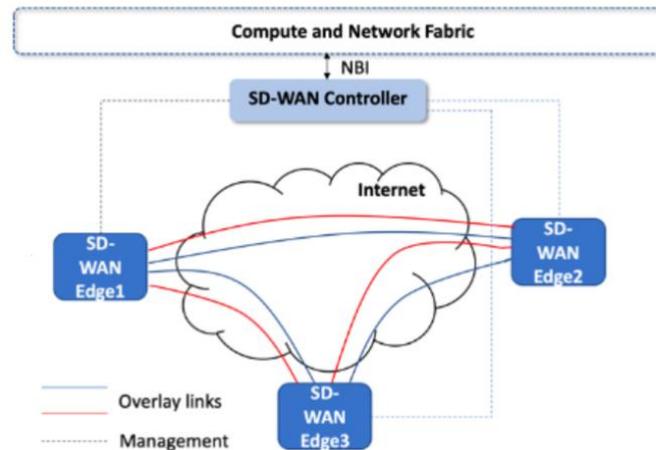
Figure 17: Network Compute Fabric and programmable SD-WAN

2.2.3.1.2.10 Potential Requirements for Use Case 1 - Scenario 2: Edge Intelligence and Compute Interconnection

Identified challenges are related to the DIMOs and Network-Compute Fabric integrated technologies, adapted to the 6G Telco Infrastructures, to provide relevant SLAs for deployed services in the envisioned different domains, intent defined instantiation and infrastructure network translated to the compute fabric, for RAN and Edge domains, on top of existing technologies within testbeds. Another challenge is related to the AI and learning capabilities of the system (MLaaS), for cognitive enablement of domains, SD-WAN interconnection with dynamic resources integration. In the end, it will also be challenging to ensure reliable operation and inter-operation of resources and services with respect to KPIs in terms of SLA, as described previously and applied to the Edge intelligence and compute interconnection Use Case experiments. From a business perspective, a new service model could emerge from the proposed PoC results, and this should be further evaluated in terms of economic/financial impact vs technological investments for a positive business plan. Other challenges of the Use Case 1 – Scenario 2 are the following:

- Validation of resource utilization efficiency for SLA, supported by unified Intent APIs (NBIs and SBIs),

- Interface towards/from tenants and Network-Compute Fabric to the SD-WAN Controller for dynamic aspects as autonomous self-configuration, self-healing, and self-optimization of envisioned network.

- Proper network integration of the SD-WAN controllers within the Network-Compute Fabric, to provide zero-touch management and configuration between edges,

- Validation of the QoS enforcement,

- Deployment of seamlessly services in the network.

The requirements for this experiment are summarized in **Table 18: Requirements for Use Case 1 Scenario 2 (Edge Intelligence and Compute Interconnection)**.

Table 18: Requirements for Use Case 1 Scenario 2 (Edge Intelligence and Compute Interconnection)

ID	UC1–S2-01
Name	CDN Service subscription
Description	Request of CDN-based video service instantiation
ID	**UC1–S2-02**
Name	Service requirements
Description	CDN-based video services optimization for the users(automated)
ID	**UC1–S2-03**
Name	Service Termination
Description	CDN-based video service termination

2.2.3.1.2.11 Radio Specific Requirements

- **Coverage**: Urban, multi-site, requires handover with minimal jitter.

- **Bandwidth**: >100Mbps, packet mode.

- **uRLLC**: Latency <10ms, Reliability >99.999%

- **Energy**: Edge sensing optimized, -30% consumption target

2.2.3.2 Use Case 2: Metaverse

The Metaverse blends physical and digital worlds into one where XR users, content, and digital entities interact. It is one of the emerging use cases, which is expected to drive the transition to 6G systems, as it requires KPI improvements by at least an order of magnitude. Extended Reality requirements are already at the limit of 5G network capabilities, requiring extremely high downlink capacity (>100Mbps per device) and low latency (<5ms) to achieve immersive operation. The Metaverse further pushes the limits, allowing users to freely navigate and interact in virtual worlds. One of the key features of the Metaverse is its fully immersive and interactive environment. Users will be able to interact with each other in real-time, which means that the network infrastructure will need to be able to handle a massive amount of data in real-time. This trend is well aligned with the ongoing digitalization of our societies and presents opportunities for telecom operators to pursue the future generation upgrade of their systems (from 5G to 6G). Recent advancements in the volumetric video domain allow to capture the user's volume in real time and insert the resulting hologram into these VEs, enabling multiple users to be co-present, together with others, in a 3D space while embodied in their own self-representation. This Use Case will demonstrate the 6G-INTENSE components' ability to support and enrich metaverse experience in the context of multi-technological domain deployment.

Use Case 2 will support two different experiments, aiming to demonstrate the ability of the DIMO components to support the Metaverse services and to improve users' experience leveraging on zero-touch configuration and management of intent-based requests, from definition to lifecycle management, using closed-control loops.

2.2.3.2.1 Use Case 2 - Scenario 1: Joint Communication and Sensing for Optimal User Tracking in the Metaverse

2.2.3.2.1.1 Description

Future 6G systems will utilize intelligent algorithms for Joint Communication and Sensing (JCS) to facilitate User Equipment (UE) localization and tracking. Integrating the ability to sense user location and movement patterns into the communication network will enable many innovative use cases, ranging from those aimed at enhancing the performance of the network itself to those targeted at offering spatial sensing capabilities as a feature to external users or applications to support pervasive location awareness. While these algorithms perform well in outdoor environments, their accuracy and precision are inadequate for indoor settings. In such scenarios, users often connect to Deep Edge devices using non-3GPP, best-effort protocols due to weak cellular signals. This connection helps offload computations and conserve battery power. In this experiment we consider a scenario of users connected through Wi-Fi to a Metaverse service deployed using the 6G-INTENSE DIMO framework. The experiment will focus on showcasing user localization and tracking capabilities based on JCS, exposed as a micro-service by a Local Manager and Orchestrator (an edge or extreme edge node). The experiment will start with devices implementing the Fine Time Measurement (FTM) protocol introduced in IEEE 802.11-2016, which enables precise indoor ranging and positioning thanks to precise time-of-flight estimation, and will evaluate the ranging precision as a function of the number of FTM-capable devices and other relevant parameters, as well as exploring the possibility to enhance the ranging through novel techniques, as the exploitation of additional information carried by the signals and extracted through Channel State Information (CSI) analysis.

The experiment will consist of two main phases:

1. Evaluating the performance of each UE-AP pair individually.

2. Evaluating the effectiveness of multilateration using FTM-derived distance estimations.

All measurements will be executed both indoors and outdoors for performance comparison.

2.2.3.2.1.2 Source

- 6G-INTENSE SNS JU web: https://smart-networks.europa.eu/phase-2-stream-b/#6G-INTENSE
- 6G-INTENSE at Europa portal: https://cordis.europa.eu/project/id/101139266
- 6G-INTENSE home WEB page: https://6g-intense.eu

2.2.3.2.1.3 Roles and Actors

- Metaverse Verticals/Application Providers: Deploy their applications or services as a 6G service on top of the 6G infrastructure and gain access to spatial sensing APIs for improved immersive interaction.

- Telecom Operators/DMOs: Leverage Deep Edge sensing to enhance service quality and energy efficiency.

- End users: Benefit from more accurate, responsive Metaverse environments.

- Infrastructure/resource providers: Deploy Wi-Fi and deep edge compute nodes for FTM processing and API exposure.

2.2.3.2.1.4 Pre-conditions

Availability of FTM-capable devices and deployed location microservices on deep edge nodes.

2.2.3.2.1.5 Triggers

User device movement, XR app request for localization, accuracy improvement tasks.

2.2.3.2.1.6 Normal Flow

1. UE connects to the Metaverse application via Wi-Fi with FTM support.

2. Access Points collect localization/mobility data.

3. JCS node processes multilateration and refines accuracy.

4. Localization/mobility data is exposed to Metaverse app via the JCS API.

5. Tracking adapts based on app energy/accuracy trade-offs.

2.2.3.2.1.7 Alternative Flow

Fallback to camera-based localization if RF sensing fails.

2.2.3.2.1.8 Post-conditions

Tracking data logged; service adapts sensing strategy for future requests.

2.2.3.2.1.9 High Level Illustration
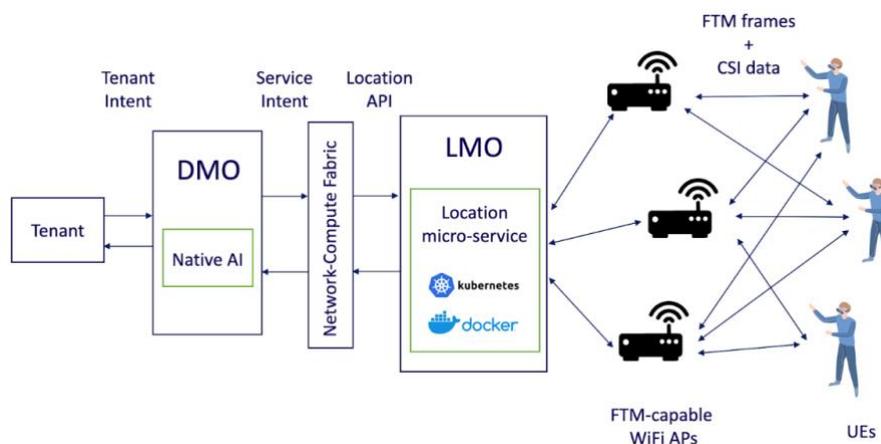
Figure 18: Architecture of Experiment 2.1

The experiment comprises the following main building blocks, as depicted in **Figure 18: Architecture of Experiment 2.1**.

- At end-user side, FTM-capable UEs are connected through non-3GPP access to a Metaverse service.

- At the deep edge, FTM-capable Wi-Fi APs perform multilateration of UEs to provide user location and tracking capabilities based on JCS. Multilateration of user devices will leverage the FTM protocol; location accuracy enhancement using Channel State Information (CSI) will be explored.

- A location micro-service, deployed at the edge on the LMO, will expose spatial sensing capabilities via Location API both to the Network-Compute Fabric and applications running on user devices.

For the initial phase of the experiment, a Google Pixel 6 smartphone with Android 13 onboard will be used as the FTM-capable UE. For the FTM-capable APs, we decided to employ two

different devices: the Aruba 505 AP and the Celer onboard automotive router, developed by Teldat. Both devices are dual radio, 5GHz and 2.4 GHz 802.11ax 2x2 MIMO. Additional devices might be considered in the future.

2.2.3.2.1.10 Potential Requirements for Use Case 2 - Scenario 1: Joint Communication and Sensing for Optimal User Tracking in the Metaverse

<u>Accuracy of FTM Depends on the Environment</u>

The accuracy of Fine Timing Measurement (FTM) is significantly influenced by the surrounding environment. One of the primary environmental factors that affect FTM accuracy is multipath propagation. In scenarios where multipath effects are prominent, the Time of Flight (ToF) measurements obtained through FTM can become increasingly noisy and unreliable. Multipath occurs when signals reflect off various surfaces before reaching the receiver, causing multiple signal paths that interfere with one another and distort the true ToF.

To mitigate these issues and restore accuracy, several strategies can be implemented:

- <u>Utilizing Channel State Information (CSI).</u> By incorporating CSI at both the APs and client devices, it is possible to better detect the time-of-first arrivals. CSI provides detailed information about the signal's propagation characteristics, enabling the system to differentiate between direct paths and reflected paths. This allows for more precise identification of the first arrival time, thereby improving the accuracy of ToF measurements.

- <u>Collecting Measurements on Multiple Channels</u>. Accuracy can be enhanced by gathering measurements across multiple channels. By dynamically switching both the client device and all APs or anchors to different channels, the system can average out the noise and errors introduced by any single channel. This multi-channel approach helps to mitigate the effects of frequency-specific interference and multipath, leading to more reliable FTM readings.

- <u>Utilizing Multiple Frequency Bands</u>. Modern wireless networks can exploit multiple frequency bands, including the traditional 2.4GHz and 5GHz bands, as well as the newer 6GHz band. Collecting measurements across these different bands can further improve accuracy. Each band has distinct propagation characteristics and multipath behaviours; combining measurements from all these bands provides a more comprehensive dataset, reducing the impact of noise and improving overall accuracy.

- <u>Exploiting Multi-Link Operations (MLO).</u> MLO enables simultaneous use of multiple frequency bands, facilitating continuous and coherent measurement collection across 2.4GHz, 5GHz, and 6GHz bands. This capability allows for better management of band-specific challenges and leverages the strengths of each band to enhance the reliability and precision of FTM measurements.

By addressing these challenges through advanced techniques such as CSI integration, multi-channel measurements, and multi-band operations, it is possible to significantly improve the accuracy and reliability of FTM in diverse and complex environments.

Accuracy of FTM Depends on the Specific Device/Chipset

The accuracy of Fine Timing Measurement (FTM) can be highly dependent on the specific device or chipset being used. Different devices and chipsets may have varying levels of precision and may introduce unique biases or errors into the measurements. This variability often necessitates a calibration step to ensure consistent and accurate results across different hardware. To address this challenge, several methods can be employed:

- Oversampling Approach: One effective method to mitigate device-specific inaccuracies is to use an oversampling approach. By increasing several Access Points (APs) or anchors involved in the measurement process, a larger dataset can be collected. This abundance of data allows for better statistical analysis, smoothing out any anomalies or errors introduced by individual devices. Additional measurements help to resolve ambiguities and compensate for any lack of calibration of client devices.

- Increasing Number of APs/Anchors: Adding more APs or anchors into the network infrastructure enhances the spatial diversity of the measurements. This increased density of measurement points helps to provide more robust data, reducing the reliance on any single device or chipset. The diverse sources of measurement can be used to cross-verify and correct the ToF data, leading to more reliable and accurate FTM results.

- Self-Measurements by Anchors: Another approach to address device-specific inaccuracies is to implement self-measurements by the anchors themselves. Since the positions and locations of the anchors are known in advance, they can perform self-calibration procedures to correct any mis-calibration or biases in their own measurements. By regularly conducting self-measurements, the anchors can maintain a high level of accuracy and ensure that their ToF data remains reliable.

- Resolving Ambiguity and Mis-Calibration: The self-measurement process helps to resolve any ambiguity or mis-calibration issues that might arise from environmental changes or hardware drift over time. Anchors can use their known positions to calibrate their internal clocks and measurement algorithms, ensuring that their ToF readings are as accurate as possible. This proactive calibration step enhances the overall reliability of the FTM system.

By employing these methods, it is possible to address the challenges posed by device and chipset variability in FTM accuracy. The combination of oversampling, increasing the number of APs/anchors, and implementing self-measurements by the anchors helps to ensure that the FTM system remains accurate and reliable, regardless of the specific hardware in use.

The requirements for this experiment are summarized in **Table 19: Requirements for Use Case 2 Scenario 1**.

Table 19: Requirements for Use Case 2 Scenario 1

ID	UC2–S1-01
Name	User localization query
Description	Services can request the location of target UEs by querying the Location API.
ID	P2–ES2.1-FR02
Name	User tracking subscription
Description	Services can subscribe to receive periodic updates about the location of target UEs.
ID	UC2–S1-02
Name	User tracking cancellation

Description	Services can cancel their tracking subscription to target UEs.
ID	**UC2–S1-03**
Name	Energy/accuracy management
Description	According to the tenant's intent, different trade-offs between location accuracy / tracking frequency and energy consumption can be enforced through the Location API. For instance, location accuracy may be lowered by decreasing the number of FTM frames that are sent by UEs for a single localization event to preserve battery power, while in tracking the tracking period can be adjusted (e.g., from 0.5s to 10s).

2.2.3.2.1.11 Radio Specific Requirements

- <u>Coverage</u>: Indoor, room-level precision, NoLOS

- <u>Bandwidth</u>: 2.4/5/6 GHz bands, >100Mbps packet mode

- <u>URLLC</u>: Tracking jitter <5ms, Reliability >99.999%

- <u>Location</u>: Accuracy <1m, battery optimized

2.2.3.2.2 Use Case 2 - Scenario 2: Fully Autonomous Metaverse FCAPS, sensing and Continuum Abstraction

2.2.3.2.2.1 Description

This scenario demonstrates end-to-end service orchestration for a Metaverse environment using the DIMO architecture, which includes intent-based deployment, service mesh packaging, and dynamic adaptation. AI-driven fault management and network/resource reconfiguration ensure SLA continuity during runtime.

We will demonstrate the DIMO components from an end-to-end point of view while showcasing intent translation, and propagation up to the infrastructure. The ORO and EUR testbeds represent the resource pools, interconnected via SD-WAN, which are onboarded via the 6G-INTENSE DMO, indicating their preferences and constraints (e.g., pricing, SLOs, Coverage) via the intent mechanism. Moreover, we consider a Metaverse Service Mesh adopting the 6G-INTENSE generalized paradigm, for the packaging of cloud-native NFs (e.g., UPF), the Metaverse Application micro-services, as well as a Cyber-Physical Systems (CPS) service for Sensing. The generalized Metaverse Service Mesh is onboarded via the DMO, with Tenants indicating their preferences and constraints in terms of Service-Level Objectives (SLOs) of individual functions, main objectives, and hard constraints with the vertical Tenant intent. The DMO intent handler loop is then responsible for the topological mapping of the Service Mesh, via an intent (re-)negotiation process between the Tenant and the Resource owner, via fully autonomous actions driven by the Native AI toolkit. Service deployment at the Southbound Resource Pools (as represented by testbeds) is facilitated by their respective Network-Compute Fabric abstraction frameworks, which have already been trained/finetuned to abstract the respective LMOs in the framework of Use Case 1. The LMOs considered are the EUR O-RAN LMO for the orchestration of the Network resources and the ICOM π-Edge for Compute Continuum orchestration on both sites. During experiment execution, users are connected to the network via O-RAN gNBs, and roaming with their UEs (e.g., in the form of VR headsets) and tracked via a Sensing service. Faults and workload variations are created to test the system's adaptability and the system's intra- and inter-domain coordination and conflict resolution among the DMO and Network-Compute Fabric, showcasing the respective HRL capabilities of the Native AI toolkit. The Use Case will show several FCAPS decisions with O-RAN and pi-Edge LMOs cooperating as an orchestration continuum, via intent adaptations, service migrations between nodes following users' movement patterns, updating network link configuration, and scaling computing resources and radio resources.

2.2.3.2.2.2 Source

- 6G-INTENSE SNS JU web: https://smart-networks.europa.eu/phase-2-stream-b/#6G-INTENSE
- 6G-INTENSE at Europa portal: https://cordis.europa.eu/project/id/101139266
- 6G-INTENSE home WEB page: https://6g-intense.eu

2.2.3.2.2.3 Roles and Actors

- <u>Vertical Tenants</u>: Declare service meshes and constraints (coverage, cost, latency).

- <u>Telecom Operators/DMOs</u>: Optimize network resource allocation across multiple domains.

- <u>AI Framework Developers</u>: Benefit from real-world validation of autonomous FCAPS.

- <u>Infrastructure/resource providers</u>: Own one or more 6G resource pools.

- <u>End users</u>: Enjoy robust, low-latency immersive experiences with transparent adaptation.

2.2.3.2.2.4 Pre-conditions

Enabled DMO and Network-Compute Fabric, trained AI models, and interlinked ORO/EUR testbeds.

2.2.3.2.2.5 Triggers

Metaverse service intent declaration, SLA degradation detection, node fault events.

2.2.3.2.2.6 Normal Flow

1. Tenant submits intent with constraints.

2. DMO creates service mesh topology and forwards the respective resource intents to Network-Compute Fabric.

3. Network-Compute Fabric abstraction layer allocates the resources via communication with the southbound platforms.

4. Native AI continuously monitors service and resource lifecycle for faults.

5. System self-heals and adapts deployment.

2.2.3.2.2.7 Alternative Flow

Fallback to default routing if dynamic optimization fails; alert tenant with updated SLA.

2.2.3.2.2.8 Post-conditions

Service re-optimized, updated AI models, logs stored for audit and retraining.

2.2.3.2.2.9 High Level Illustration

Same architecture as in Scenario 1 of Use Case 2 shall be used.

2.2.3.2.2.10 Potential Requirements for Use Case 2 - Scenario 2: Fully Autonomous Metaverse FCAPS, sensing and Continuum Abstraction

The challenges of this Use Case are related to the DIMO architecture and its main components, namely, the Network Compute Fabric and DMO. The first challenge will be regarding the intent definition, an intent model should be designed for different levels of architecture (business, service, resource). Native AI machine learning models should be trained, especially models that operate in the intent translation loop, these models should be trained and tested against typical intents that can be received. Models related to FCAPS should also be trained to respond to the introduced faults (network faults, compute resources faults, node faults ...). Additionally, it will be challenging to ensure reliable operation and inter-operation of resources and services with respect to KPIs in terms of SLA, as described previously and applied to the Edge intelligence and compute interconnection Use Case experiments.

The requirements for this experiment are summarized in **Table 20: Requirements for Use Case 2 Scenario 2**.

Table 20: Requirements for Use Case 2 Scenario 2

ID	UC2–S2-01
Name	Metaverse Service Mesh onboarding
Description	A Metaverse Service Mesh is onboarded at multiple resource pools via 6G-INTENSE DMO; preferences and constraints (e.g., pricing, SLOs, Coverage) expressed via intent mechanism
ID	UC2–S2-02
Name	Metaverse Service Mesh instantiation
Description	The DMO intent handler loop is responsible for the topological mapping of the Service Mesh, via an intent (re-)negotiation process between the Tenant and the Resource owner, via fully autonomous actions driven by the Native AI toolkit
ID	UC2–S2-03
Name	User Tracking
Description	During experiment execution, users are connected to the network via O-RAN gNBs, and roaming with their UEs (e.g., in the form of VR headsets) and tracked via a Sensing service.
ID	UC2–S2-04
Name	Automated Fault Management
Description	Faults and workload variations are created to test the system's adaptability and the system's intra- and inter- domain coordination and conflict resolution

2.2.3.2.2.11 Radio Specific Requirements

- <u>Coverage</u>: Cross-domain, requires seamless handover

2.2.3.2.2.12 Bandwidth:

- >100Mbps packet-based

2.2.3.2.2.13 URLLC

- Latency <5ms,
- Resilience to dynamic conditions

2.2.3.2.2.14 Other requirements

- <u>Power</u>: Adaptive management, with wearable/VR headset constraints

2.2.4 SUNRISE-6G (SUstainable federatioN of Research Infrastructures for Scaling-up Experimentation in 6G) Use Cases

"*Network of Networks*" concept of 6G in the SUNRISE-6G project

6G is expected to emerge as key enabler for the intelligent digital society of 2030 and beyond, providing superior performance via ground-breaking access technologies, such as joint communication and sensing, cell-free, Radio Intelligent Surfaces, and ubiquitous wireless intelligence. Most importantly, 6G is expected to trigger a total rethink of network architecture design, which builds on the key idea of new stakeholders entering value chain of future networks. The SUNRISE-6G (https://cordis.europa.eu/project/id/101139257) approach is inspired by the "network of networks" concept of 6G Networks, aiming to integrate all private and public infrastructures under a massively scalable internet-like architecture. SUNRISE-6G aspires to create a federation of 6G test infrastructures in a pan-European facility that will support converged Testing as a Service (TaaS) workflows and tools, a unified catalogue of 6G enablers publicly accessible by experimenters, and cross-domain vertical application onboarding.

The business roles below, common for all use cases, have been defined based on identified relations:

- **Application Business Owner (ABO)**: stakeholder responsible for setting the UC within the experimentation plane. Acts as a mediator between end users and all other testbed-related members of business relations organises and manages services to be provided by the testbed, and does not act as a direct consumer.

- **Application User (AU)**: vertical end users (consumers/prosumers). This group is the project's targeted audience, which resides at the forefront of the value chain. SUNRISE-6G is expected to bring real value to end users with a complete thing-to-cloud value chain. It will provide new experiences with new innovative services and applications. The public will better understand how novel services that demand beyond today's 5G performance characteristics can be used to provide better experiences for future- services with guaranteed Quality of Experience (QoE).

- **Testbed Business Owner (TBO):** a stakeholder capable of leasing testbed resources on a commercial basis, separated from maintenance or specific operations. Acts as a testbed service provider, confirming Service-Level Agreements (SLAs) for testbed services can be achieved.

- **Testbed Feature Contributor (TFC):** Testbed facility stakeholders or external contractors mainly focus on extending testbeds' technical features/frameworks. This role expands the testbed's capabilities, thus widening the variety of experiments it can execute.

- **Testbed Operator (TO)**: stakeholder responsible for carrying out tests and maintaining the testbed. Fulfils SLA agreements, define restrictions for possible tests on the testbed, and ensure its availability.

- **Testbed Federation Provider/Integrator (TFPI):** testbed facility stakeholder or external contractor responsible for unifying APIs/VPN access and deploying federation components to integrate a testbed into a larger pool of testbeds as an equal resource.

- **Network/Cloud Provider (NCP):** testbed facility stakeholder or external contractor providing network connectivity or cloud storage services.

- **Technology Provider (TP):** external contractor offering services for experimentation plane UCs.

- **Equipment Vendor/Supplier (EVS):** testbed facility stakeholder or external contractor providing hardware components such as compute units/devices or network controllers.

- **Standardisation (S):** stakeholder with active participation in various standardisation organisations, contributing to new standards, bringing in new requirements to identify and fill gaps, and evolving existing standards related to 6G architectures.

- **Researcher (R):** stakeholder focusing on research and development activities.

<u>**Summary of SUNRISE-6G Use Cases**</u>

The following table presents a high-level overview of the four use cases evaluated within the SUNRISE-6G framework. Each use case explores distinct domains, ranging from immersive XR to industrial automation and NTN service continuity, highlighting the project's diverse technological applicability across vertical sectors. The summary includes primary focus, enabling technologies, and key stakeholders for each scenario, offering a reference for standardization bodies, industrial adopters, and research communities.

No.	Use Case Title	Primary Focus	Key Technologies	Main Stakeholders
UC1	Federated Metaverse	Real-time XR services with federated 6G testbeds	CAMARA APIs Edge Cloud, Federation XR HoloMIT	I2CAT, KUL XR Service Providers
UC2	Collaborative Robotic Exploration of Mining Environments	Autonomous robotic mining with edge offloading	Edge AI LiDAR streaming private 6G SOP Federation	LTU-RAI, UOULU Robotics/Mining Industry
UC3	Mobile Industrial Assets aided by NTN Edge Cloud & Communications	Service continuity for mobile logistics assets via NTN	NTN Edge SOP handover satellite + terrestrial access	Logistics Operators NTN Providers SOP Integrators

2.2.4.1 Use Case 1 : Federated Metaverse

2.2.4.1.1 Description

Despite major advances and increased maturity, the eXtended Reality (XR) ecosystem still faces key performance, interoperability, sustainability, and cost barriers when targeting ubiquitous networked services over heterogeneous environments. While such challenges are not limited to any single specific technological pillar, next-generation cloud continuum and network technologies can become fundamental enablers to overcoming current limitations in the XR domain.

Figure 19: Federated Metaverse UC1 visual

Besides increased ubiquity and performance (mainly in reliability, latency, and bandwidth), 6G cellular networks allow for adaptable configurations and expose resources and functions programmatically and elastically to target services. This results in emerging service-oriented paradigms and even associated business models termed NaaS, in which networks are turned into application platforms. The advent of 6G networks results in a sustainability trend where MNOs can share resources, creating a comprehensive cloud continuum through federation. In this context, SUNRISE-6G aims to provide cross-domain 6G Federation Platforms powered with innovative enablers, like Cognitive Cloud Continuum, network exposure and aggregation capabilities and associated NaaS APIs. Such contributions will offer remarkable benefits to XR services, and they will be assessed and demonstrated for a Federated Metaverse UC, which will support real-time multiuser holographic communications with users connected across cross-country environments via inter-domain networks.

The objective of this Federated Metaverse UC is to enable ubiquitous, adaptive, and scalable XR holographic communication services with realistic and volumetric user representations captured in real-time and integrated into a shared Metaverse scenario when the involved users are connected through remote and federate (cross-country) 6G experimentation facilities, by leveraging key 6G enablers from SUNRISE-6G (Federation Framework, NaaS APIs, Cognitive Cloud Continuum and Orchestration policies…). Most state-of-the-art multiuser Social XR or Metaverse platforms currently use synthetic avatars to represent users. However, recent studies have reflected on the remarkable benefits of providing realistic user representations in Metaverse-related services, such as trustworthiness, levels of embodiment, co-presence, and quality of interaction.

The SUNRISE-6G counts with a fully functional and worldwide pioneering Metaverse platform called HoloMIT, able to provide real-time multiuser holographic communications with realistic and volumetric user (self and others) representations, using low-cost and off-the-shelf XR capture and rendering hardware (**Figure 20**). Up to date, HoloMIT has been assessed and validated in different scenarios for a set of relevant UCs, such as collaborative viewing, interactive virtual events, and virtual meetings, obtaining satisfactory and promising results. However, state-of-the-art holographic communication (and Metaverse-like) technologies and platforms still encounter key needs and challenges in terms of mainly:

- performance (e.g., delays, fps, resolution, stability…)

- resources consumption (processing, bandwidth…)

- adaptability and scalability (and thus inter-operability)

- effective deployment over multi-domain and heterogeneous scenarios

The SUNRISE-6G will provide new 6G enablers to enhance HoloMIT versatility and performance, allowing for its successful deployment and usage in ubiquitous scenarios.

Figure 20: Examples of holographic communication scenarios using the HoloMIT platform

HoloMIT has been built by integrating modular client-based (e.g., multi-sensor captures sub-systems, encoding/decoding, presentation, interaction features) and server-based (e.g., Session and Media Managers) components and modules, which are deployed in traditional client-server architectures, either with usage of Selective Forwarding Units (SFU) or of Multipoint Control Units (MCU), with well-defined interfaces. The SUNRISE-6G aims at exploiting the modularity of HoloMIT and the complete control over its end-to-end components to extend it with novel interfaces with 6G enablers from the project that will:

- devising and adopting an elastic and scalable architecture with the assistance of Cloud Continuum and Federation capabilities;

- adapting transversal XR media functions (e.g., forwarding, transcoding…) for their strategic orchestration and deployment over the Cloud Continuum, based on needs/demands;

- devising and adopting strategic mechanisms to boost uplink/downlink capacity by aggregating carriers and transport layer connections

- leveraging and adopting Network Exposure Functions (NEF) and associated service-level APIs to notify about congestion situations and/or request QoD features.

Use Case 1 - Scenario 1: Federation via Edge-Cloud capabilities and APIs

The SUNRISE-6G will offer APIs to identify Edge Cloud resources, where Virtualised Network Functions (VNF) for XR processing and communication functions (e.g., Selective Forwarding Units (SFUs) and even transcoding units) can be instantiated in a strategic, flexible, and agile manner. Diverse Edge selection criteria can be configured based on location, available resources, or business-related aspects (e.g., costs or agreements with specific providers / Operators). Additionally, these Edge Cloud APIs can provide capabilities to facilitate VNF/AF lifecycle management. The first phase of this experimentation case aims to demonstrate the feasibility of running holographic communication sessions across federated domains and associated performance levels. The second phase will explore the benefits of the Federation Framework and associated Edge-Cloud APIs.

Use Case 1 - Scenario 2: *Exposure and Consumption of Network Capabilities*

The SUNRISE-6G will expose advanced network capabilities via simplified APIs. These include Rate Recommendations based on detected congestion (or other network analytics/QoS metrics) and the ability to request QoD to be enforced via diverse network capabilities, such as QoS Class identifier (QCI) prioritisation – to boost the priority of one of the flows; Aggregate Maximum Bit-Rate (AMBR) limitation – to protect the flow of interest by limiting the max rate of other flows; and Access Traffic Steering, Switching and Splitting (ATSSS), to aggregate, e.g. 5G and Wi-Fi segments, if available. Such experimental cases will require the extension of both User Plane XR AFs—to perform the required media adaptations—and Control Plane XR AFs to appropriately interface with the NaaS API via both NBI and potentially EWBI (in case of federation).

2.2.4.1.2 Source

- SUNRISE-6G European SNS JU Project – https://cordis.europa.eu/project/id/101139257

2.2.4.1.3 Roles and Actors

Actors:

Business roles (and provided benefits)	Partners
Application Business Owner (ABO) **(They are provided with network and infrastructure resources and federation platforms, so their applications/services can be deployed and offered over ubiquitous cross-country and cross-operator environments)**	I2CAT
Testbed Business Owner (TBO) **(They can monetise leasing of resources and exposure of capabilities/NaaS APIs)**	I2CAT KUL
Network/Cloud Provider (NCP) **(They can monetise Edge-Cloud APIs so that their Edge resources can be discovered and selected by ABO)**	I2CAT KUL

- XR Users and Developers (APs)

- Federation Managers and SOP Integrators

- Edge Resource Providers

- Orchestration Platform Administrators

2.2.4.1.3 Pre-conditions

Deployment of HoloMIT XR platform shall be done at two federated testbeds (e.g., I2CAT and KUL). Federation framework must be operational, including SOP, EWBI, and CAMARA NBI APIs. Edge Cloud nodes with sufficient compute resources will be available and discoverable via NaaS APIs. The HoloMIT platform shall be integrated and compatible with CAMARA APIs for dynamic service deployment across edge/cloud nodes.

2.2.4.1.4 Triggers

Initiation of a multi-user XR session requiring real-time holographic communication. Detection of network congestion or user request for improved QoS/QoE also triggers rate adaptation or resource reassignment via APIs. User requests for initiating multi-party XR sessions or detection of conditions requiring bandwidth adaptation or offloading (e.g., network congestion, user join/leave events, or location-triggered events).

2.2.4.1.5 Normal Flow

- (1) XR application (HoloMIT) requests resource discovery using CAMARA Edge APIs.

- (2) SOP's Open Exposure Gateway queries local and federated zones via Federation Manager.

- (3) Application metadata and artefacts are onboarded through the SOP to selected zones.

- (4) Application instantiation proceeds via SRM and orchestration triggers deployment on edge nodes.

- (5) NaaS APIs allow runtime adaptations e.g. bandwidth prioritisation, QoS provision, or ATSSS activation.

- (6) Performance metrics and telemetry are gathered via SOP interfaces and exposed to the AP.

2.2.4.1.6 Alternative Flow

If the primary testbed does not have sufficient resources, the Federation Manager will redirect the deployment request to a federated SOP. If the federation fails, fallback to degraded XR service using pre-cached low-fidelity assets is applied. Rate limitation and adaptive encoding is performed dynamically to maintain user QoE.

2.2.4.1.7 Post-conditions

Application session ends and teardown commands are issued via SOP. Resources are de-allocated, performance logs stored, and any gathered metrics can be published for research or SLA evaluation. Successful or partial completion of XR sessions. KPIs, usage logs, and QoE metrics are stored. Resources are released and the federation registry is updated. Optional dataset anonymization and publication may follow for standardization or research.

2.2.4.1.8 High Level Illustration
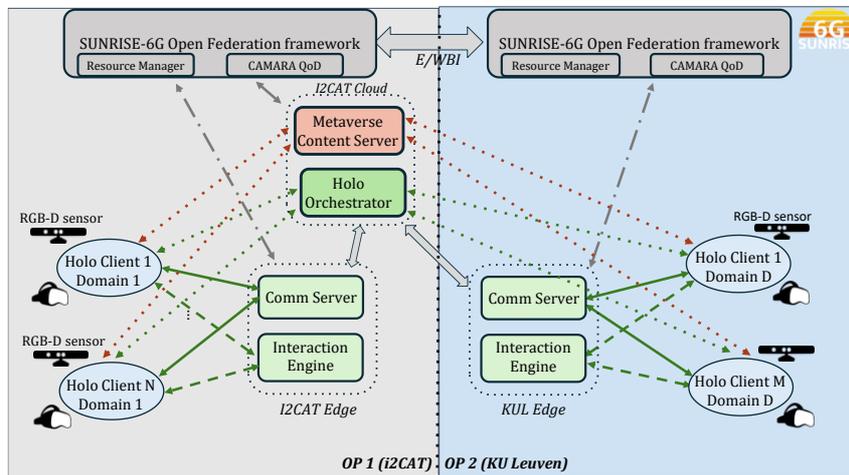

Figure 21: Envisioned evolution of HoloMIT in SUNRISE-6G[3]

The **Figure 21** sketches the envisioned platform evolution to contribute to the next-generation affordable and interoperable holographic communication services with increased performance, adaptability, scalability, ubiquity, and user satisfaction.

---

[3] boxes with dotted red lines represent components to be evolved

This will also contribute to higher sustainability (due to resource usage and cost minimisation and the availability of an effective medium reducing the need to commute for face-to-face gatherings). Cross-domain federation and network exposure capabilities will be leveraged and demonstrated in the Metaverse UC, by adopting and adapting HoloMIT, which will be deployed at 2 SUNRISE-6G testbeds (i.e., I2CAT and KUL) that will represent two Metaverse rooms. CAMARA APIs will be leveraged to seamlessly deploy cloud/edge HoloMIT services at any testbed node (cloud or edge) completely unchanged.

The components include:

- XR Clients (users with capture/render devices)

- SOP components (OEG, AEM, FM, SRM)

- Edge/Cloud Nodes

- Federation interfaces (E/WBIs) linking SOPs at I2CAT and KUL

- Data flows for registration, deployment, and adaptation.

2.2.4.1.10 Potential Requirements

Functional Requirements:

Rate Adaptability: the XR services need to be able to dynamically adapt the transmitted video quality/bitrate to adhere to the available needs and/or resources based on recommendations provided by Control Plane network elements

Real-time cross-layer KPI metrics measurement and monitoring: the capability to measure, report, monitor, and potentially expose, in real-time, a wide variety of cross-layer KPIs (e.g., performance and resource usage levels) from the XR service.

Edge Computing and Orchestration: Capabilities to leverage Edge Computing & Orchestration enablers to offload XR processing functions from clients or to bring communication servers to most appropriate Edge.

Federation Capabilities: The 6G network shall offer APIs to identify Edge Cloud resources, where VNF for XR processing and communication functions can be instantiated in a strategic, flexible, and agile manner.

Exposure Capabilities: The 6G network shall expose advanced network capabilities via simplified APIs to the XR services, like Rate recommendations based on detected congestion (or other network analytics/QoS metrics), and the ability to request QoD to be enforced via diverse network capabilities, such as: QCI prioritisation – to boost the priority of one of flows; AMBR limitation – to protect the flow of interest by limiting max rate of other flows; and ATSSS – to aggregate 5G & Wi-Fi segments. Such features to be provided through Northbound Interface (NBI) to Control Plane XR Application Functions (AF).

Non-functional Requirements:

End-to-end latency: the end-to-end latency for XR communications over federated domains shall be lower than 200ms.

Scalability: sessions with up to 8 concurrent users shall be supported stably.

Resources Saving: Edge Computing and NaaS APIs capabilities shall allow the saving of processing resources up to 30% and bandwidth up to 50%, with a focus on the client side.

Increased Quality of Experience (QoE): Edge Computing and NaaS API capabilities shall result in increased QoE (at least 1 point on the Mean Opinion Score (MOS) scale) for the end-users.

2.2.4.1.11 Radio Specific requirements

Radio Coverage

- Coverage across indoor and outdoor XR deployment zones at I2CAT and KUL.

- Handover: Multicell and seamless handover required.

- Mobility support for users walking within room-scale XR range (≤5 km/h).

- Expected typical range: 100–300 meters (urban/indoor LoS)

- Coverage: Crosses public and semi-public indoor spaces

- Multicell required: YES – scope: city, campus

- Seamless handover required; delay tolerance: <50 ms

- Mobility: pedestrian (~5 km/h) and seated users

- No maritime or aerial coverage required

2.2.4.1.12 Bandwidth requirements

- Peak uplink/downlink up to 100 Mbps per user; average 30–50 Mbps.

- Packet mode traffic with variable bitrate encoding.

- Peak data rate: 100 Mbps per user (upstream)

- Average data rate: 50 Mbps

- Traffic: Packet-switched, video and control streams

2.2.4.1.13 URLLC requirements

- Required latency: < 200 ms (one way)

- Required reliability: 99.999% during active session transmission

- Max jitter: < 20 ms

2.2.4.1.14 Radio regimens requirements

- Preferred: Licensed public mobile spectrum (5G/6G NR)

- Acceptable: Shared licensed or managed unlicensed bands (e.g., Wi-Fi 6E) supported under ATSSS.

2.2.4.1.15 Other requirements

- Location services required with accuracy <1m for indoor positioning

- UE power consumption: moderate to high

- Battery: Rechargeable, minimum 2 hours for continuous capture/render

Business Requirements
Relevance for at least 3 Business Roles from the XR Application, Testbeds and Network/Cloud Providers, with >=2 KPIs and >=2 KVIs associated with each of them (Use SUNRISE-6G to ensure the set of applications operates correctly and meets its intended functionality)
<u>**Open Access Datasets:**</u> **Provide valuable Open-Access datasets (KPIs, KVIs, XR traffic patterns…, etc.) for the scientific and standardisation communities**

2.2.4.2 Use Case 2: Collaborative Robotic Exploration of Mining Environments

2.2.4.2.1 Description

Robotic applications, particularly in the mining industry, increasingly rely on autonomous multi-agent systems for monitoring and exploration tasks. Nowadays, most developed solutions often use centralised and onboard control algorithms based on peer-to-peer communications and, more recently, on the advanced features of 5G networks. The latter allows for partially offloading computationally demanding tasks from the onboard platform to 5G-enabled edge cloud servers. Although this helps with compute-intensive workloads, 5G-edge capabilities still introduce unwanted latency and are insufficient for multi-agent high-capacity uplink missions, making them unsuitable for time-sensitive, computation-heavy tasks.

Figure 22: Robotic Exploration of Mining Environments UC2 visual

The 6G-enabled edge computing is expected to provide a solution by offering low-latency, scalable, and self-healing capabilities. Robots will communicate and cooperate in future mining environments via edge-based architectures, enabling real-time decision-making and task execution. Centralised schemes will optimise system performance, while edge technologies manage computational demands across clusters. This approach is critical for the autonomous machines of tomorrow.

The UC2 will take place in the Sub-Terranean (Sub-T) laboratory of the Robotics and AI (RAI) group at LTU. The laboratory is an underground facility consisting of multiple tunnels of varying lengths, sizes, and shapes, uniquely suited for validating and demonstrating robotic applications for the mining industry. A picture of the main tunnel is visible in **Figure 23.**
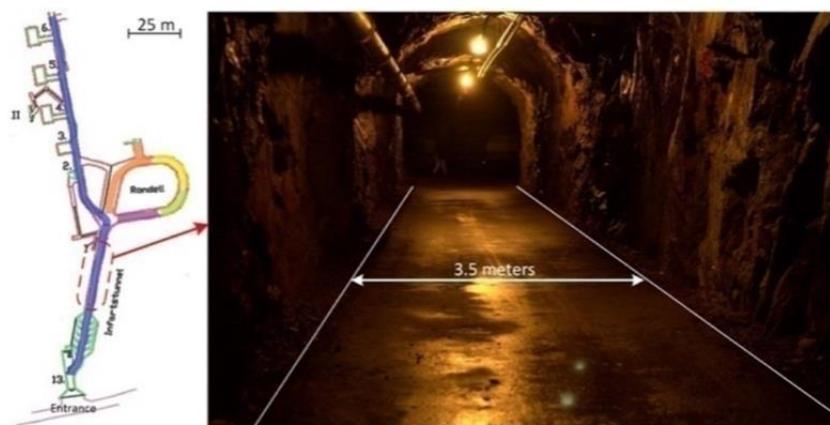
Figure 23: Main tunnel and partial map of the Sub-Terranean laboratory of the RAI group

UOULU will provide a 6G private network within the Sub-T tunnel environment to enable this scenario. To showcase the federation of private networks, the UOULU facility will implement the Slices Blueprint federation model and connect the private 6G network installed in the Sub-T Lab to the SUNRISE-6G network through it. On the other hand, the RAI group at LTU will utilise the 6G network for advanced computing offloading tasks via 6G-enabled edge cloud servers. This will reduce the computational requirements on the onboard robotics platforms for complex missions and enable the deployment of multi-agent systems. These systems will involve multiple robotic machines offloading tasks and communicating with each other for advanced swarming operations (e.g., mining exploration and mapping). Additional 6G features, such as localisation, adaptive QoS, and AI capabilities, will also be explored. In the reference test scenario, multiple components are involved. More in detail, it is possible to identify four main components:

1. The RAI's robotic platforms are connected to the 6G network in the Sub-T lab.

2. The 6G network testbed deployed in the Sub-T lab, allowing low latency and high throughput communication between the robotic platforms and the 6G-edge cloud server.

3. The OULU or LTU 6G-edge cloud server, used to offload computing intensive tasks, multi-agent robot's controllers, etc., and the optional local breakout of the OULU network located in the LTU facility. Capabilities can also be delivered through a "network on wheels" type system instead of the local breakout.

4. Additional components of the 6G network belonging to the network federation enable advanced 6G-framework capabilities (e.g., E2EAI framework, etc.).

2.2.4.2.2 Source

- SUNRISE-6G European SNS JU Project – https://cordis.europa.eu/project/id/101139257

2.2.4.2.3 Roles and Actors

Business roles	Partners
Subterranean Testbed Facility Provider	LTU-RAI
6G Network Provider	UOULU
Network Federation Integrator	UOULU
Robotics Platform Provider	LTU-RAI
System Integration and Deployment	LTU-RAI, UOULU

Business roles	Partners
Open Dataset Creation and Management	LTU-RAI, UOULU
Edge Computing Infrastructure Provider	LTU-RAI
Robots Controllers and Data Processing Algorithms Developments	LTU-RAI

Actors:

- Autonomous Robots (UEs)

- Edge Cloud Server

- Local 6G Core Network and UPF

- Open Federator Interfaces between LTU and UOULU

2.2.4.2.4 Pre-conditions

Deployment of a 6G private network in the Sub-T lab at LTU, including edge computing servers, connected robotic platforms, and integration with the SUNRISE-6G federation via OULU federated core components. Robot agents must be configured for uplink streaming and control via the edge cloud.

2.2.4.2.5 Triggers

Initiation of exploration or mapping mission requiring computational offload, real-time coordination, or advanced swarm control. Tasks like LiDAR streaming, radio coverage scanning, or multi-agent cooperation may trigger edge/cloud resource requests.

2.2.4.2.6 Normal Flow

1. Robotic agents connect to the 6G RAN in the Sub-T Lab.

2. Robots begin data streaming to the edge server via UPF (edge breakout).

3. Edge server performs offloaded LiDAR/image data processing and decision-making logic.

4. Robots coordinate in real-time based on edge-derived insights.

5. Network metrics and robot telemetry are monitored to trigger adaptive QoS or load balancing.

6. If needed, control plane triggers federation APIs to connect remote compute/network functions.

7. Data and results (e.g., maps, datasets) are logged and optionally shared.

2.2.4.2.7 Alternative Flow

If local edge resources are insufficient or unavailable, computation tasks may be redirected to the federated edge via OULU's network using the Open Federator framework. If connectivity drops, fallback onboard algorithms ensure robots perform minimal safe tasks autonomously.

2.2.4.2.8 Post-conditions

Robot missions are completed with relevant results collected. Raw and processed data (LiDAR, maps, radio KPIs) are stored locally and optionally shared as open datasets. The network returns to standby or low-power mode until next trigger.

2.2.4.2.9 High Level Illustration

The **Figure 24**'s schematic representation provides a high-level overview of the components involved in the mentioned UC. This includes an Edge node and RAN deployment delivered by the Luleå Technical University and a 5G Core located in the Oulu 5G Test Network. The SUNRISE-6G Open Federator framework is leveraged to interconnect the LTU deployment with the UPF of the 5G Core in Oulu. Using UPF, data from the robots is routed directly to the Edge server, which is available locally in the Sub-T Lab.
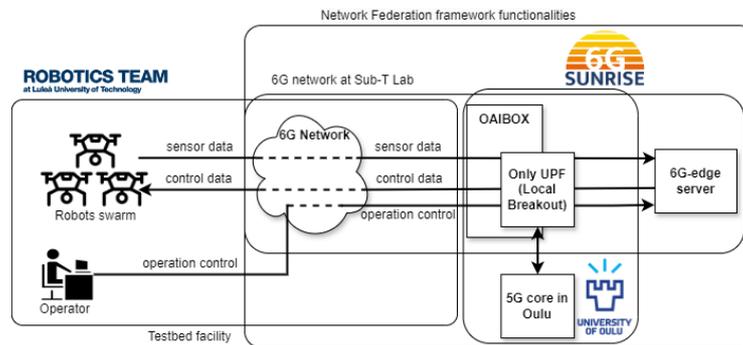
Figure 24: Robotic 6G-enabled offloading for mining environments use case

This UC scenario leverages 6G-enabled platforms to showcase multi-agent missions for exploration and mapping tasks (e.g., LiDAR and radio coverage mapping in the tunnels), as well as the usage of a 6G-edge system for advanced computational sensor dataflows processing offloading (e.g., offload LiDAR and/or high-quality image streams processing in the 6G-edge). Based on the available network capabilities (in terms of available throughput, latency between the robots and the edge server, as well as signal quality and coverage in the tunnels of the Sub-T testbed facility), the robotic platforms will try to exploit the advanced offloading capabilities of the 6G-edge server, thus enabling complex real-time robotics operations in mining-like environments. The UC will try to reach and test the boundaries of the 6G network to optimise the robotics applications suitable for mining operations. SUNRISE-6G technologies involved in the UC2 are E2EAI and Open Federator frameworks. The open dataset to be produced in the experiments is called Dataset on Three-Dimensional (3D) Point Cloud scans of a subterranean area.

The diagram includes the following components:

- Multi-agent robotic platforms

- Sub-T Lab 6G network and edge node

- OULU testbed and federation link

- Edge/cloud compute and storage units

- Data/control paths for streaming, processing, and coordination.

2.2.4.2.10 Potential Requirements

Functional Requirements:

<u>Low-latency communication:</u> the network must provide one-way latency below 10 ms between the UE and the Edge server, enabling real-time autonomous robot operations.

<u>High Uplink Capacity:</u> the network must support uplink throughput of up to 100 Mbps to handle data-intensive tasks like LiDAR and high-quality image streaming from multiple agents.

<u>Edge Computing Offloading:</u> enable dynamic offloading of computation-heavy tasks (e.g., LiDAR/image processing) to the edge server to ensure onboard computational efficiency.

<u>Multi-Agent Control:</u> the system must support multi-agent coordination and communication protocols that enable collaborative tasks like mapping and exploration.

<u>Adaptive QoS Prioritization:</u> implement adaptive network QoS to allocate network resources dynamically between different UEs based on task and device priorities.

<u>Federation with SUNRISE-6G Network:</u> ensure seamless integration of Sub-T Lab's 6G network with the broader SUNRISE-6G network federation.

Non-functional Requirements:

<u>Architecture Scalability:</u> the architecture should support scaling regarding the number of robots, sensor data, and computational load without degrading performance.

<u>Data Security and Privacy:</u> secure all communications and data transmissions, especially inter-network communications within the SUNRISE-6G network federation, to ensure compliance with privacy and security standards.

<u>Real-Time Performance Monitoring:</u> continuous monitoring of the relevant network KPIs, like latency, throughput, and the main radio quality indexes, to maintain real-time operation standards.

2.2.4.2.11 Radio Specific requirements

Radio Coverage

- <u>Radio range</u>: tens to hundreds of meters indoors (Sub-T tunnels)

- <u>LOS</u>: Primarily non-line-of-sight (NoLOS)

- <u>Multicell</u>: YES – scope: facility/building

- Seamless handover: desirable

- <u>Mobility</u>: slow to moderate (~0–10 km/h)

- <u>Special needs</u>: subterranean coverage

2.2.4.2.12 Bandwidth requirements

- <u>Peak data rate</u>: 100 Mbps (uplink per robot)
- <u>Average</u>: 50 Mbps
- <u>Traffic</u>: packet-switched video, telemetry, LiDAR
- <u>Isochronicity</u> not required

2.2.4.2.13 URLLC requirements

- <u>Latency</u>: <10 ms one-way

- <u>Reliability</u>: ≥99.999%

- <u>Jitter</u>: <5 ms

2.2.4.2.14 Radio regimens requirements

- <u>Preferred</u>: licensed spectrum (private 6G network)

- <u>Acceptable</u>: research spectrum for trial deployments

2.2.4.2.15 Other requirements

- Low power consumption per UE; must support continuous operation for hours

- Rechargeable battery systems

- Terminal location needed with sub-meter accuracy in 3D space

Business Requirements
<u>**Partner Contributions Management**</u>: **Identify each partner's roles and contributions to enable effective collaboration within the defined UC.**
<u>**Open Dataset Availability**</u>: **Establish protocols for creating and sharing the dataset combining radio KPIs and 3D Point Cloud scans of a Sub-T lab as a public resource for future development.**

2.2.4.3 Use Case 3: Mobile Industrial Assets aided by NTN Edge Cloud & Communications

2.2.4.3.1 Description

This use case aims to support mobile industrial operations—such as trains, buses, or cargo carriers—by enabling service continuity through Non-Terrestrial Network (NTN) Edge Cloud connectivity. The use case focuses on maintaining low-latency, high-reliability access to compute and network services for connected assets moving through rural or remote regions. By leveraging SUNRISE-6G federated platforms and open APIs, the system ensures that industrial assets maintain service performance during connectivity transitions, including satellite-based access.

Figure 25: NTN in 6G UC3 experiments 1 and 2 visuals

Efficient public infrastructures, such as those facilitating product logistics and people transportation, are crucial for fostering a sustainable society, economy, and growth. These infrastructures typically encompass highly distributed systems' operation and automation, integrating mobile and fixed assets. Examples are sectors like railways, public transport, and logistics. To ensure continuous operation, infrastructure solutions and services must be resilient and reliable in urban and rural settings, which heavily depends on robust hardware, connectivity, and computational resources. However, regarding reliable hardware, standalone building becomes extremely costly, whereas resource sharing frequently impacts reliability.

Non-terrestrial (such as space-based) resources appear as a good candidate to tackle that challenge. These systems offer a unique advantage by being independent of local (terrestrial) disturbances and can be used ad hoc to provide coverage around the globe. This capability is particularly beneficial for operating mobility infrastructures and assets, ensuring seamless functionality regardless of geographic location.

The key objective of this UC is to assess and demonstrate the benefits that NTNs can introduce on the end-2-end (E2E) operation of industrial services. It will focus on individual experiments that validate the following benefits:

1. **Continuity and reliability:** Implement and demonstrate orchestration and management mechanisms for an operation designed and maintained to provide continuous operation over an extended period with minimal interruption.

2. **Resilience:** Demonstrate scenarios that withstand and recover from disruptions and leverage the dynamics and flexibility of communication and computing infrastructure.

3. **Efficiency:** implement mechanisms to optimally and sustain the use of limited resources.

4. **Safety:** more reliable and resilient systems significantly enhance safety by ensuring continuous operation and reducing the risk of failures.

These scenarios will provide insights into how future 6G systems can leverage satellite-terrestrial convergence to support diverse and challenging UCs, thus leading to more resilient, flexible, and ubiquitous connectivity solutions. This UC will use the NTN & NTCS Testbed from Siemens for experimentation given its capabilities to emulate/simulate non-terrestrial edge cloud resources and satellite communication links. The testbed can emulate real world satellite systems/clusters and allows to create custom topologies that are envisioned in the next generation of satellite constellations. Moreover, the UC will consider existing HW/SW components to be integrated in the experiments, given the possibility of integrating real HW/SW components. The UC scenarios will be implemented according to Supervisory Control and Data Acquisition (SCADA) concepts. Industrial SCADA systems are very common in today's management and operation of critical infrastructure across domains, including railways, public transport, and logistics:

- In railways, SCADA systems are used to monitor and control train movements, track conditions, and signalling systems, ensuring safe and efficient operations. They provide real-time data on train locations, speeds, and schedules, allowing for quick responses to any disruptions or emergencies.

- In public transport, SCADA systems help manage bus and tram networks by monitoring vehicle locations, passenger loads, and traffic conditions. This enables operators to optimise routes, reduce delays, and improve overall service reliability.

- For logistics, SCADA systems are crucial in managing supply chain operations, including warehouse management, inventory control, and transportation logistics. They provide visibility into the movement of goods, helping to streamline operations, reduce costs, and enhance delivery times.

One of the scenarios will be selected and recreated in the UC as an example of a critical infrastructure application. The application would be later used to validate the 5-fold benefits introduced by NTN: continuity, reliability, resilience, efficiency, and safety.

Top-Level Design

A high-level overview of the components involved in the UC is represented at **Figure 26.**
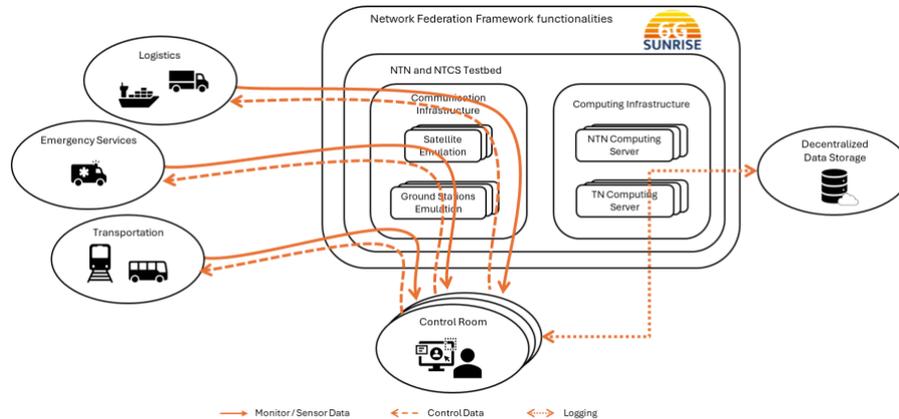
Figure 26: Mobile Industrial Assets supported by NTN Edge Cloud and Communications UC

In a reference test scenario, multiple components are involved, namely:

1. Simplified field devices - sensors and actuators corresponding to the lowest level of a SCADA system.

2. The NTN & NTCS Testbed supporting:

- Emulated satellite infrastructure to enable the NTN-based communication between satellite-enabled entities.

- Space-based Edge/Clouds infrastructure available as part of the computing payloads of the emulated satellite infrastructure.

3. Simplified control room: Supervisory computers collect data, which is then used to monitor and control the process.

4. Decentralized data storage to keep all historical data, which can then be used for data analysis, reporting and/or generating alerts based on the collected data.

All components, along with an implemented scenario application, will enable 5-fold benefits testing.

Use Case 3 - Scenario 1: Multi-RAT over 5G NR and legacy NTN

Mobile and Static UEs are equipped with multiple and heterogenous radio access technologies (RAT), providing connectivity via Terrestrial 5G and legacy Non-Terrestrial Networks (respectively, 5G NR and Satellite connectivity). This dual-connectivity capability is used to simultaneously provide redundant paths from the field devices to the control-centre. The level at which redundant connection is created can either be at the network level (i.e., IP/TCP) or at application-level. The experiment to be tested is described as follows: due to the mobility of mobile assets, one of the radio links is lost. However, existing sessions are maintained another disjoint path on from a different radio technology still in-place.

Use Case 3 - Scenario 2: *Satellite Edges for Decentralized Data Storages*

Satellite links are limited, and multiple connections of the same UE over the satellite link are inefficient. Thus, using both the uplink and downlink to exchange huge volumes of data might become inefficient and costly. Decentralised data storage, deployed at the Satellite Edges, can enable efficient balance between data transmissions and availability. The experiment to be tested is described as follows: decentralised data storage will be deployed as part of the Satellite Edge infrastructure, and access will impact availability, efficiency, and data transmission.

Use Case 3 - Scenario 3: *Migration of application context* *(no details are currently available)*

As the satellites move in their orbit, the application context of the satellite goes through a period of getting far away - getting closer - getting far away. This will likely create high oscillations in the E2E latency and a waste of Inter-Satellite Link (ISL) resources. The experiment to be tested is described as follows: application context (such as session information) will be migrated to satellites near the field devices so that the data is always kept close to the source (or destination).

2.2.4.3.2 Source

- SUNRISE-6G European SNS JU Project – https://cordis.europa.eu/project/id/101139257

2.2.4.3.3 Roles and Actors

Roles (Use Case 3- Scenario 1):

Business roles	Partners
Application Business Owner **(Use SUNRISE-6G to ensure the set of applications operates correctly and meets its intended functionality)**	SIE
Application User **(Test the set of applications across a representative environment for an E2E deployment offered by SUNRISE-6G and validate against the service KPIs of the target deployments)**	SIE
Testbed Operator **(Create measurement data based on the application user's service descriptions and running services)**	SIE
Testbed Business Owner **(Define, maintain, and govern standardised interfaces from the Testbed Federation Provider)**	SIE

Roles (Use Case 3 - Scenario 2):

Business roles	Partners
Testbed Feature Contributor **(Use SUNRISE-6G to ensure that the developed features of the platform operate correctly and achieve their intended functionality)**	INFO
Researcher **(Test various scenarios, enabled by the developed features, across a representative environment offered by SUNRISE-6G and validate the corresponding KPIs)**	INFO NCSRD OTE
Testbed Operator **(Create measurement data based on the application user's service descriptions and running services)**	NCSRD OTE

Actors:

- Industrial trains, logistics carriers

- Edge compute nodes (terrestrial and airborne)

- CAMARA-compliant platform exposing APIs

- SOP with FM and SRM managing dynamic connectivity and compute exposure

2.2.4.3.4 Pre-conditions

NTN connectivity infrastructure must be available and integrated with edge cloud resources. Mobile industrial assets must be equipped with suitable communication modules to switch between terrestrial and non-terrestrial links. Federated SOPs with exposed CAMARA APIs must be in place to support application deployment and continuity.

2.2.4.3.5 Triggers

Asset movement into an area with limited terrestrial coverage. Onboard system requests resource exposure via CAMARA APIs or initiates application migration. Detected QoS degradation may also trigger edge failover or service redeployment.

2.2.4.3.6 Normal Flow

1. The mobile asset connects to the closest available terrestrial edge node.

2. A running application consumes compute services and telemetry is transmitted.

3. As coverage changes, the FM monitors network status, and triggers service migration.

4. Application artifacts are shared between SOPs via EWBI.

5. The application is re-instantiated at an NTN edge node.

6. The mobile asset switches to satellite uplink and resumes operations.

7. All transitions are monitored for SLA conformance and QoS metrics.

2.2.4.3.7 Alternative Flow

If no suitable NTN edge is available, the system retains minimal client-side fallback logic. Tasks that are latency-sensitive are temporarily buffered or run with reduced fidelity until connectivity resumes.

2.2.4.3.8 Post-conditions

Application continues execution with possible minimal degradation. Metrics collected from both terrestrial and NTN execution phases are stored and used for SLA compliance verification and dataset creation.

2.2.4.3.9 High Level Illustration

Entities:

- Industrial vehicle

- terrestrial edge

- NTN edge node

- CAMARA API exposure

- SOP with FM/SRM

Interactions:

- data flow for monitoring

- API queries for edge capabilities

- application instantiation workflows

For architectural details refer to section 0 above.

2.2.4.3.10 Potential Requirements

Functional Requirements (Use Case 3 - Scenario 1):

Redundancy support (i.e., terrestrial and NTN connectivity). The system must support using different and disjoint network paths so that the applications can switch between them or use them simultaneously.

Satellite infrastructure shall support Satellite Computing, including placing service components in satellites that support custom edge computing (MEC). This requires interfaces to deploy and manage services, containers, or VMs in satellites.

Handovers must be fast, with low losses and transparent for the applications.

Functional Requirements (Use Case 3 - Scenario 2):

Dual Connectivity Support: The system must support dual connectivity, allowing simultaneous connections to terrestrial and satellite networks.

Inter-Network Handover Support: The system must support handovers between terrestrial and non-terrestrial (satellite) networks, specifically between different PLMNs.

Session Continuity: The system must maintain active sessions without interruption during handover and support seamless traffic balancing across both networks.

MPTCP Integration: The system must support MPTCP, enabling multiple sub-flows from both terrestrial and satellite networks within a single TCP session.

Non-functional Requirements (Use Case 3 - Scenario 1):

Availability of 99.99% and 99% for critical and non-critical services.

Reliability of 99.99% and 99% for critical and non-critical services.

Saving of at least 30% for data transmissions in the satellite's downlink.

High security that follows zero trust principles.

Non-functional Requirements (Use Case 3 - Scenario 2):

Latency Tolerance: The system should maintain end-to-end latency within an acceptable range for high-priority services during handover and dual connectivity usage.

Throughput Optimization: The system should ensure high throughput performance during dual connectivity, leveraging MPTCP to maximize available bandwidth.

Seamless User Experience: User experience should remain seamless during network handover, with minimal perceived service disruption.

Fault Tolerance: The system must be fault-tolerant, with failover mechanisms to maintain connectivity even if one network (terrestrial or satellite) fails.

System Monitoring and Logging: The system must include robust logging and monitoring features to track handover and connectivity performance, assisting in troubleshooting and performance analysis.

2.2.4.3.11 Radio Specific requirements

Radio Coverage

- <u>Typical range</u>: kilometres via NTN link

- <u>LOS</u> is ideal for NTN, terrestrial fallback where available

- <u>Multicell</u>: YES – across wide regional scope

- <u>Handover</u>: required between satellite and terrestrial; tolerable impact <1 s

- <u>Mobility</u>: high (train, vehicle speed up to 300 km/h)

2.2.4.3.12 Bandwidth requirements

- <u>Peak</u>: 100 Mbps downlink, 50 Mbps uplink

- <u>Average</u>: 20–50 Mbps

- <u>Traffic</u>: Packet-switched, adaptive streaming and telemetry

2.2.4.3.13 URLLC requirements

- <u>Latency</u>: 10–100 ms depending on link (lower for terrestrial)

- <u>Reliability</u>: 99.999% required for safety and operations

- <u>Jitter</u>: <20 ms acceptable

2.2.4.3.14 Radio regimens requirements

- Licensed satellite bands (e.g., S/L/Ka)

- Public mobile 5G licensed fallback in terrestrial coverage areas

2.2.4.3.15 Other requirements

Other requirements (Use Case 3 - Scenario 1):

Requirement	Description	Rationale
Standardised Interfaces	This requirement pertains to defining consistent interfaces for interacting with the testbeds. These interfaces ensure uniformity and ease of integration.	Standardised interfaces facilitate interoperability, allowing different subsystems to communicate seamlessly.
Standardised Service Descriptors	This requirement involves documenting detailed descriptions of the services provided by the Application Users / Application Business Owner. It includes information on functionality, behaviour, and usage.	Clear service descriptions help users understand what each service offers and how to utilise it effectively.

Requirement	Description	Rationale
Isolation of Experiments	Ensure that experiments run in isolated environments, preserve data privacy, and prevent interference between concurrent experiments.	Isolation enhances security, prevents unintended side effects, and maintains the integrity of experimental results.
Reliable Experiment Scheduling	Implementing a reliable scheduling mechanism for experiments, allowing users to plan and execute tests at specific times.	Scheduled experiments enable systematic testing and feedback collection.
Monitoring of Experiment and Infrastructure	Collecting measurement and status data from the testbed infrastructure. This includes subsystem performance, resource utilisation, and overall health.	Monitoring ensures proper functioning, identifies issues, and supports troubleshooting.
Quality Information of the Testbed	This provides information about the testbed's quality, capabilities, and limitations so users can assess whether it meets their application evaluation needs.	Transparency helps users make informed decisions.
Metadata for Measurement Data Sets	Describing the structure and context of measurement data sets generated during experiments. Metadata includes timestamps, units, and relevant parameters.	Metadata ensures data traceability, reproducibility, and meaningful analysis.
Monitoring Data at Different Levels	Collecting monitoring data at various levels, such as infrastructure, service, and end-to-end. This includes performance metrics, logs, and alerts.	Comprehensive monitoring aids in diagnosing issues and optimising system performance.
Infrastructure Support for Reproducibility	Designing the infrastructure to support the reproducibility of test runs. This involves preserving experiment configurations, input data, and execution context.	Reproducibility ensures that others can validate and build upon experimental results.

Other requirements (Use Case 3 - Scenario 2):

Requirement	Description	Rationale
Standardised Interfaces	This requirement pertains to defining consistent interfaces for interacting with the testbeds. These interfaces ensure uniformity and ease of integration.	Standardised interfaces facilitate interoperability, allowing different subsystems to communicate seamlessly.
Standardised Service Descriptors	This requirement involves documenting detailed descriptions of the services provided by the Application Users / Application Business Owner. It includes information on functionality, behaviour, and usage.	Clear service descriptions help users understand what each service offers and how to utilise it effectively.
Isolation of Experiments	Ensure that experiments run in isolated environments, preserve data privacy, and prevent interference between concurrent experiments.	Isolation enhances security, prevents unintended side effects, and maintains the integrity of experimental results.
Monitoring of Experiment and Infrastructure	Collecting measurement and status data from the testbed infrastructure. This includes subsystem performance, resource utilisation, and overall health.	Monitoring ensures proper functioning, identifies issues, and supports troubleshooting.
Metadata for Measurement Data Sets	Describing the structure and context of measurement data sets generated during experiments. Metadata includes timestamps, units, and relevant parameters.	Metadata ensures data traceability, reproducibility, and meaningful analysis.

Common business requirements for all scenarios:

Business Requirements
Relevance for at least 3 Business Roles from the public infrastructure, transportation, and logistics.   **(Use SUNRISE-6G to ensure the set of applications operates correctly and meets its intended functionality)**
<u>**Open Datasets Availability**</u>: Provision of Open Access Datasets as a valuable resource for R&D and scientific communities.

2.3   Drones

2.3.1   An innovative fire detection pilot solution using 5G, Artificial Intelligence and drone technology

An innovative fire detection pilot solution using 5G, Artificial Intelligence and drone technology

2.3.1.1 Description

Provide motivation of having this use case, e.g., is it currently applied and successful; What are the business drivers, e.g., several stakeholder types will participate and profit from this use case

Provide on a high level, the operation of the use case, i.e., which sequence of steps are used in this operation?

Wildfires represent a significant natural risk causing economic losses, human death and environmental damage. In recent years, the world has seen an increase in fire intensity and frequency. Research has been conducted towards the development of dedicated solutions for wildland fire assistance and fighting. Systems were proposed for the remote detection and tracking of fires. These systems have shown improvements in the area of efficient data collection and fire characterization within small-scale environments. However, wildland fires cover large areas making some of the proposed ground-based systems unsuitable for optimal coverage.

To tackle this limitation, unmanned aerial vehicles (UAV) and unmanned aerial systems (UAS) were proposed along with ground sensors. The Sensors which are installed in strategic points in the park, are interconnected with the incident management platform and the drone control system. The drone operates scheduled surveillance flights as well as emergency flights in case of the sensor indications.

The system is able to detect smoke or fire, both by the sensors indications at the field and from specific algorithms that are used to analyse drones' video in real-time. In both cases the data are send to the Control Center indicating points of interest.

When a sensor identifies abnormal values of CO_2 or/and temperature sends an alarm to the Control Center with the coordinate of the event. At that point two actions take place:

An SMS / Email is sent to the involved stakeholders with the exact location of the event

The drone autonomously takes-off and is directed straight ahead to the indicated location to verify the event with the help of the Ai algorithms. The drone during all operations broadcasts live to all stakeholders that are involved.

In the case of a preprogramed patrolling where the drone detects smoke or fire through the camera, it sends an alert to the control centre, and the drone, immediately rushes to where the smoke was detected to verify the incident and send the exact location info. Then the drone either returns to its base or records the progression of the fire. The result is the immediate identification of the starting point of the fire, real-time monitoring of remote areas, early visual detection of smoke and fire, and in result protection of human life.

2.3.1.2 Source

2.3.1.3 Roles and Actors

Citizens & Vicinity. People who lives (near) a critical infrastructure and needs to be protected or informed about potential risk that could affect their lives.

Critical Infrastructure. Central element source of vulnerabilities that can become real risks (natural or cyber risks).

Emergency Bodies. Stakeholders dedicated to minimizing the effects of the risks once them happens (hospitals, fireman's, etc.).

Governmental bodies. Stakeholders required to organize the society and provide insights at higher level.

Civil Protection Organization. Stakeholders dedicated to mobilizing and organize the citizens in emergency situations.

2.3.1.4 Pre-conditions

The main pre-condition here is the occurrence of an extreme event, such as a fire, that would result in severe social, environmental, and economic impacts.

2.3.1.5 Triggers

The triggers used in this use-case is when an extreme event is detected early enough in the critical infrastructure.

2.3.1.6 Normal Flow

What is the normal flow of exchanged data between the key entities used in this use case: devices, IoT platform, infrastructure, pedestrians, vehicles, etc.

The main goals are to provide a reliable early warning system in case of extreme environmental events. A prerequisite is the interoperability of the system and the data it produces with smart city standards, and the effective integration of legacy third-party applications and IoT subsystems and equipment already installed in cities.

The service aims to:

1) Surveillance, in real time, of large areas presenting a high level of risk and an increased possibility of fire through a network of ground sensors and UAV/Drone.

2) Immediate smoke or fire detection in 2 ways:

   a) Ground sensors: temperature, smoke, etc.

   b) On-board sensors in UAVs/drones (optical cameras, thermal cameras, sniffers) Optional

3) Timely confirmation of an outbreak using special high-end small-sized drones, equipped with a special camera, operating in the designated area.

4) Upon a true confirmation immediate alerts and notifications with emergency bodies (in case of required) and civil protection bodies.

5) Provide highly accurate information about the location, spread, and intensity of fires, allowing emergency responders to make informed decisions about how to respond

6) Continuous surveillance and data collection during the fire event and after. The resulting data are kept in a file (log files) and are available for further statistical analysis, patterns identification, etc. for the creation of forecasts and operational models for more efficient management of the phenomena.

7) Develop a holistic platform to provide Common Operational Picture (COP) with critical information to help decision-makers prioritize resources and respond more effectively

reducing the damage caused by fires. Reduce the need for large-scale firefighting operations and the costs associated with them.

2.3.1.7 Alternative Flow

None

2.3.1.8 Post-conditions

Continuous surveillance and data collection during the fire event and after. The resulting data are kept in a file (log files) and are available for further statistical analysis, patterns identification, etc. for the creation of forecasts and operational models for more efficient management of the phenomena.

2.3.1.9 High Level Illustration
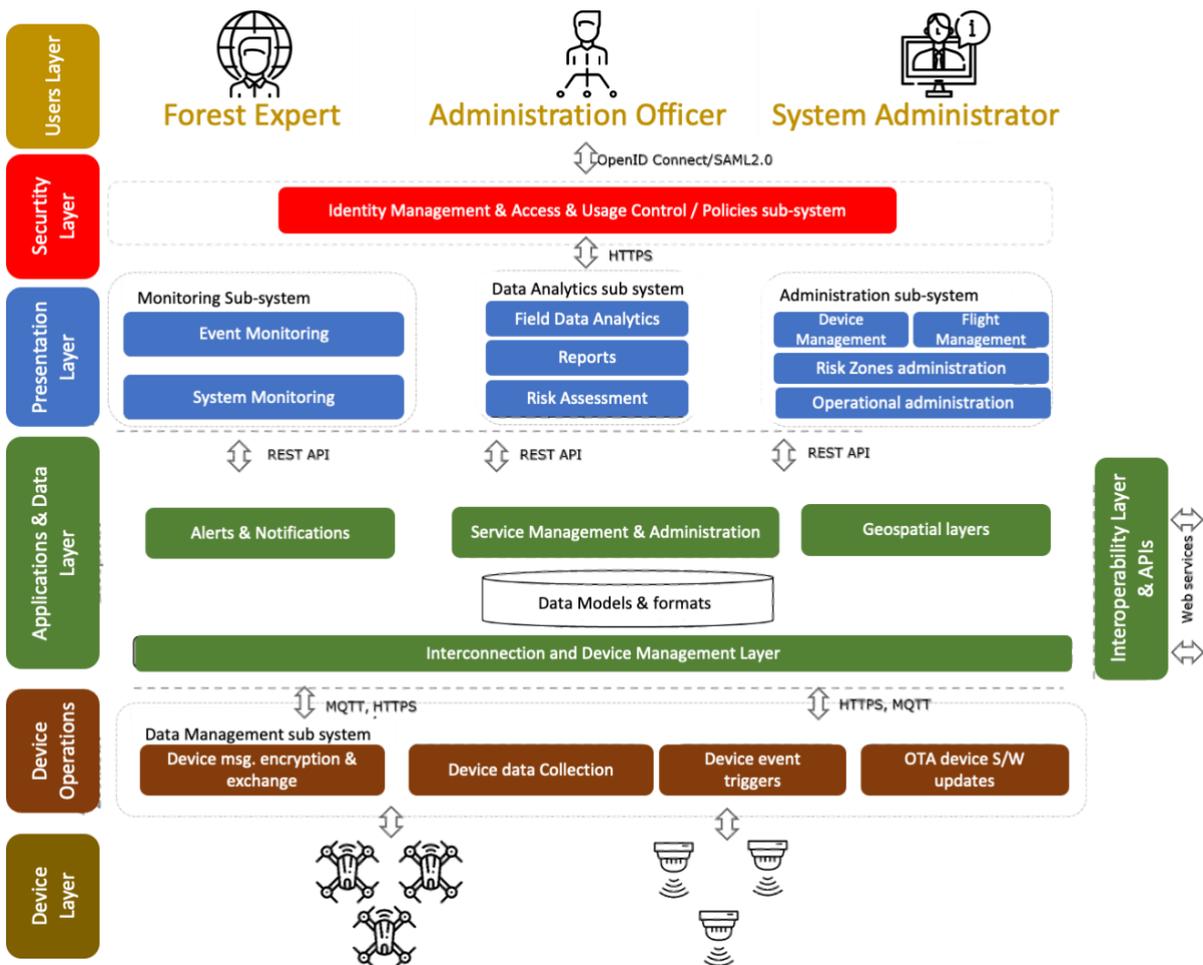
2.3.1.10 Potential Requirements

Functional Requirements

Real-time communication with the stakeholders in case of emergency.

Reliable communication between the stakeholders.

Scalable communication between systems to interconnects different critical infrastructures.

Standard-based communication between critical infrastructure to align emergency information exchange with new and legacy systems.

Non-Functional Requirements.

Secure communication between the emergency bodies due to the information nature.

Interoperability between communication protocols (linked also with the possibility to use standard communication protocols between the systems).

2.3.1.11 Radio Specific requirements

2.3.1.11.1 Radio Coverage

According to [MuBo22]:

"When working with a UAV, it is essential to control and receive image and video data remotely. Therefore, the line-of-sight, 4G/LTE, and SATCOM communication methods were used to secure the capability of operating under various circumstances and the UAV operation at long distances from the ground control station due to the size of the forest area.

A typical transmission structure contains a line-of-sight ground control station using a radio connection. It includes two datalinks (the primary one, used for image and video and telemetry exchange within 180+ kilometre range, and the backup one, for telemetry only), with automatic hopping between them in case of Global Navigation Satellite System (GNSS) or signals loss and advanced encryption standard AES-256 encryption. Secure VPN technologies, including TLS, IPSec, L2T, and PPTP, are used for data transport. This method allows the ground control station to connect with the UAV regardless of range restrictions and provide reliable cellular service. The modem concurrently enrols itself in the networks of two distinct cellular network operators and then chooses the most reliable one. Line-of-sight communications have some disadvantages, considering the range and the possibility of weather interference. SATCOM has historically been considered a Beyond Line of Sight (BLOS) communication system that would guarantee a constant connection and reliable data transmission at predetermined distances. A highly directed L-band antenna ensures a small radio signature. Furthermore, it complies with BRENT, STU-IIIB, TACLANE, STE, and KIV-7 are only some of the encryption and secure communication standards. AI server computer is located in the ground control station to process received image and video data from UAVs", copied from [MuBo22].

Moreover, according to [SiBa23]:

"To achieve secure and reliable communication for drones using a cellular communication system, drones have to exchange the information with the pilot, nearby other drones or UAVs, and principally with the air traffic control system. This mechanism is called UAV Control and Non-payload Communication (CNPC) simultaneously, depending upon the applications, a drone has to transmit or receive information on a timely basis related to the assigned task, such that images, videos, and data packets from ground entities to the drone and vice-versa. This operation is known as payload communication. To de- ploy the UAVs application on a large scale the International Telecommunication Union (ITU) has categorized the CNPC in the following section:

1. UAV Command and Control Communication (C2): This type of communication includes UAV or drone's status, a real-time control signal from pilot to UAV, and flight command updates.

2. Air Traffic Control (ATC) Relay Communication: Communication between the air traffic control system and UAV operator via ATC relay.

3. Communication for Detect and Avoid Collision: Capability to sense and avoid collision from nearby UAVs and territory.

Payload communication and CNPC require different set of spectrum. Table 2 and table 3 represents the network key points for UAV's communication. These communication parameters are specified in Release 17 by the 3GPP standards.

UAV Control and Non-payload Communication: **Table 21** represents the required QoS parameters for the CNPC communication. Here, uplink (UL) data transmission represents UAV to network side messages and downlink (DL) data transmission represents network to UAV side messages. Control and command communication is duplex communication and it may be integrated with video for controlling the operation of UAVs. Therefore, when a C2 message is sent with video, the required end-to-end latency is 1 second. A positive acknowledgment message for downlink transmission is necessary in this mode. On the other hand, when a C2 message is sent without video, end-to-end latency would be less than 40 milliseconds. This mode also requires a positive acknowledgment in downlink transmission. To communicate with the ATC relay, end-to-end latency should not be more than 5 seconds. To sense and avoid the collision with other UAVs and territories, the delay for the uplink transmission should be less than 140 milliseconds and in downlink transmission required delay is 10 milliseconds. In this mode, the reliability of the network should be 99.99% for the uplink transmission and 99% for the downlink transmission.", copied from [SiBa23].

Table 21: UAV control and non-payload communication requirements, copied from [SiBa23]

Control and non-payload communication	Message interval (UP/DL)	Message size (UP/DL)(byte)	Max UAV speed (km/h)	End-to-end latency (UP/DL)	Reliability (UP/DL)	ACK (UP/DL)
Control & Command message (without video)	1 s/ >= 1 s	84–140/100	300	1 s/1 s	99.9%	Not required/Required
Control & Command message (With Video)	40 ms/40 ms	84–120/24	60	40 ms/40 ms	99.9%	Not required/Required
Communication with UTM or ATC	1 s/1 s	1500/10K	300	5 s/5 s	99.9%	Required/Required
Detect & Avoid collision with other UAV	500 ms/500 ms	4K/4K	50	140 ms/10 ms	99.99%/99%	Required/Required

2.3.1.12 Bandwidth requirements

According to [SiBa23]:

"UAV Payload Communication: The 5G cellular technology shall be capable to transmit data collected by the entity which are installed on UAVs, such as a camera to transmit images, videos, and data files. Depending upon the applications, UAVs require different uplink and downlink quality of service (QoS). **Table 22** introduces the UAV payload communication requirements.

Table 23 introduces the communication requirements from Drone based applications.

Table 22: UAV payload communication Requirements, copied from **[SiBa23]**

UAV applications	Above ground level (m)	Max UAV speed (km/h)	End-to-end latency (UP/DL)(ms)	Data Rate (UP/DL)
8K Video Real-Time Broadcasting	<100	60	200/20	100 Mbps/600 kbps
4X4K AI Surveillance	<200	60	20/20	120 Mbps/50 Mbps
Remote UAV Controller Through HD Video	<300	160	100/20	25 Mbps/300 kbps

To transmit real-time video using a UAV up to 100 meters above ground level requires a 100 Mbps data rate for uplink transmission and 600 Kbps for downlink transmission. The allowed latency is 200 and 20 milliseconds for uplink and downlink transmission respectively. Using a UAV for surveillance needs 20 milliseconds of end-to-end latency in both uplink and downlink transmission. The essential data rate for this kind of application is 120 Mbps for uplink and 50 Mbps for downlink transmission. For controlling an UAV through HD video where the speed of the UAV is less than 160 km/h, the required uplink data rate is 25 Mbps and the downlink data rate is 20 Mbps. For this kind of application, end-to-end latency is 100 and 20 milliseconds for uplink and downlink transmission, respectively.", copied [SiBa23].

Table 23: Communication requirements from Drone based applications, *copied from* [SiBa23]

Drone based application sector	Coverage height (m)	End-to-end latency (ms)	Throughput requirements (UL/DL)
Delivery of goods	100	500	200 kbps/300 kbps
Videography and image capturing	100	500	30 Mbps/300 kbps
Security and inspection	100	3000	10 Mbps/300 kbps
Drone fleet show	200	100	200 kbps/200 kbps
Agriculture	300	500	200 kbps/300 kbps
Rescue mission	100	500	6 Mbps/300 kbps

2.3.1.13 Other requirements

Unmanned aerial vehicles, or drones, are to become an integral part of the equipment used by firefighters to monitor wildfires. They shall be used as autonomous and manual intervention remotely operated sensing platforms with AI for fire detection prevention, providing real time connectivity in a control centre. In such a holistic approach the following requirement shall be addressed

UAV types

Specialized fire **surveillance UAVs**, capable of flying in harsh weather conditions of wind, rain, extreme heat or cold, equipped with a camera that can zoom and detect fires on the fly, with an automatic health and battery status check system. The UAVs are intended for patrolling and surveillance of specific danger zones, which will be determined by the risk analysis and fire protection study.

Specialized **small confirmation drone quadcopters** for immediacy and operational risk reduction with high-end thermal and optical camera, capable of flying near high temperatures, waterproof, with automatic health and battery status check that will aim to confirm an incident on the ground.

A specialized **medium-sized UAV** that allows for ad-hoc flights on a case-by-case request basis, which should have a high-end thermal and optical camera and automatic health and battery status checks. This UAV has two (2) operational roles:

- Monitoring for smoke and fires
- Event confirmation from a local sensor or surveillance UAV

Drone Charging/Landing-Take-off Bases

The aim is for the drones to be constantly within the geographical area they are expected to operate so that they are always 'ready' to flight and thus reducing response time required. These bases must necessarily be equipped with a meteorological station that collects data in real time such as humidity, temperature, wind speed, etc. These indications must be visible both from the operations centre and from the pilots. The pilots and the operations centre, in consultation with the flight controller are taking into account all the parameters (meteorological data, flight restrictions of the drone), in order to decide whether or not the flight can be carried out. Thus, all data that the pilots process with the flight controller contribute to the commissioning or de-commissioning of the flights. Such data are recorded in a data storage kept in the operations centre. Data can be sent via 3G/4G/5G and/or WiFi with PC support on the base.

For the proper and uninterrupted operation of the bases, a charging power supply unit (UPS) capable of meeting the requirements for continuous power supply for at least 8 hours is mandatory.

Unmanned Aircraft System (UAS)

The information system consists of autonomous functional units (subsystems) that complete the infrastructure and communicate through well-defined standards and interfaces (APIs). Such subsystems of the system are:

Drone/UAV flight and control unit

Take-off/landing and charging base monitoring unit

Weather update unit

Civil aviation aircraft and drone/UAV air traffic information unit

Infrastructure orchestration and cloud interoperability extension module

2.4   Edge-Cloud Orchestration

The main goal of CODECO is to research, implement and validate a novel cognitive, cross-layer and highly adaptive Edge-Cloud management framework, which will enable flexible and effective orchestration of decentralized data workflows, dynamic offloading of computation and adaptive networking services across the Cloud/Edge computing continuum. CODECO will be validated in the scope of **six innovative use cases** that are destined to showcase the value-added Cloud/Edge functionalities of the CODECO framework, including functionalities like latency and power efficiency optimization, real-time computation adjustments, as well as flexible and adaptive networking infrastructures from the far Edge to the Cloud. The use cases are aimed at demonstrating the whole range of CODECO functionalities and features in a wide array of deployment configurations serving the needs of different stakeholders like infrastructure providers and Cloud/Edge application developers.

The initial CODECO containerized application orchestration framework composes of modular micro-services illustrated in **Figure 27**, to support the following aspects:

Automated configuration, related with application setup and application runtime across Edge-Cloud, by taking into consideration compute, network, and data aspects. Automated configuration is handled by the CODECO *Advanced Configuration and Management (ACM)* component.

Data as a resource. CODECO addresses, via its *Metadata Manager (MDM)* component data as a resource in the sense that available snapshots from the overall Edge-Cloud infrastructure, integrating different perspectives (application, user, system, data, network) at different instants of the CODECO operational workflow can be provided to different CODECO components, to assist in detecting relevant changes.

Dynamic scheduling and workload migration is supported by the CODECO component *Scheduling and Workload Migration (SWM)*. SWM integrates a novel concept by Siemens, seamless computing, including a novel scheduler for Kubernetes (K8s) that considers data-network-computation requirements to provide a best match between application requirements and available infrastructure (nodes, their computational and data properties, as well as network nodes and links), and to schedule and re-schedule application workloads across single cluster and federated cluster environments, considering application and user requirements.

Context-awareness and privacy preserving decentralised learning, supported in the component *Privacy-preserving Decentralised Learning (PDLC)*. CODECO relies on context-awareness to be able to achieve a joint data-network-compute orchestration, and on privacy-preserving decentralised learning and inference to best support readjustment of aspects such as the processing capability, computational resources, networking resources and interconnections in real-time.

Infrastructure adaptation based on a cross-layer data-compute-network approach. Via the CODECO *Network Management and Adaptation component (NetMA)*, CODECO assists in adapting not just computational (node resources) but also the networking infrastructure interconnecting such nodes.

CODECO as a framework shall support the setup of applications across clusters (the so-called K8s application deployment) and the cluster runtime management operations for single and multi-cluster environments. Users relying on CODECO during an application deployment are named as **user DEV** in CODECO. Users relying on CODECO during the cluster runtime management are referred to as **user MGT.**

Components that are expected, at the current initial stage of development (M6) to be co-located with the K8s control plane are ACM and SWM. All the other components are expected to reside in both worker nodes and if required, they may also operate on the K8s control plane.
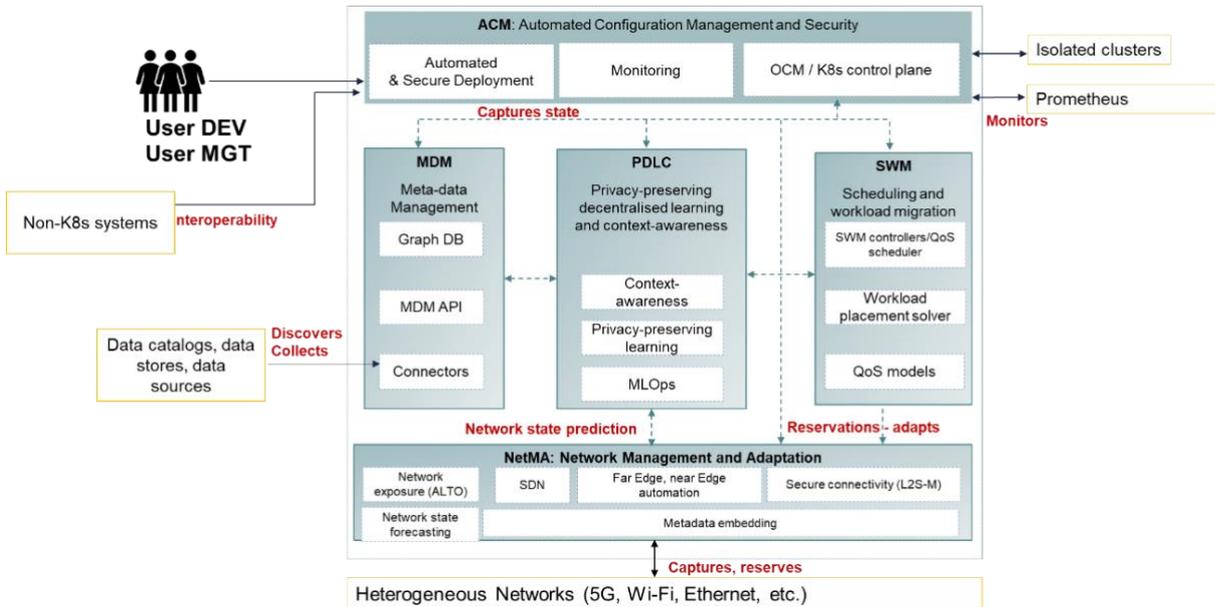

Figure 27: The CODECO K8s framework and its components

2.4.1   CODECO P1: Smart Monitoring of the Public Infrastructure

Contact: University of Göttingen, Tingting Yuan (tingting.yuan@cs.uni-goettingen.de), Xiaoming Fu (fu@cs.uni-goettingen.de)

2.4.1.1 Description

In recent years, there has been a remarkable increase in the use of Smart City solutions aimed at benefiting both residents and visitors. Those usually are achieved by leveraging advanced technology and data analytics to monitor traffic/ detect pedestrians, etc. The data collected from various sensors and devices is transmitted and analysed to optimize traffic patterns/ city planning and improve the quality of life for everyone. Now, with CODECO we can achieve a more connected, efficient, and sustainable solution. CODECO addresses the demand for handling large amounts of data in smart cities, including those with low latency demands. It is capable of orchestrating data flow across diverse features, both in terms of computation and networking. Besides, CODECO ensures a smooth and secure integration of data across Edge-Cloud environments. This ensures that data flows seamlessly and securely from Edge devices to the Cloud and vice versa, enabling end-to-end integration of smart city services. CODECO's decentralized approach to orchestration, along with its ability to deploy services in isolated, self-sufficient containers, offers great flexibility and adaptability. This means that smart city services can be deployed and executed in any environment, and the networking infrastructure can adapt to the needs of the services and the surrounding environment. This helps in creating a more resilient and adaptable smart city infrastructure.

The global purpose of P1 is to improve traffic flow and pedestrian safety in the city of Göttingen and assist in strengthening the existing Smart City concept through the implementation of a road monitoring and analytics system at the far Edge. This system comprises two parts: traffic monitoring at the city periphery, and pedestrian distribution monitoring in the city centre.

By collecting and analysing at the Edge valuable data on traffic and pedestrian behaviour, this use case aims to optimize management, reduce congestion, and enhance overall pedestrian safety and comfort, while also providing valuable insights for city planning.

In the initial phase of the pilot, P1 will focus on two specific zones in Göttingen: the city periphery, which experiences high volumes of vehicular traffic, and the city centre, where pedestrian activity is most concentrated. On a first phase of operation, these two areas shall be considered to integrate a single cluster (together with the Cloud server(s) operated by the city and UGOE). On a second phase, the two areas shall be configured as two distinct clusters.

The periphery of the city will be equipped with a combination of thermal cameras, computing units, and communication units. This will enable the real-time collection and analysis of traffic data, tracking vehicle counts and congestion levels. These insights will be used to optimize traffic flow, reduce bottlenecks, and improve overall traffic efficiency. In parallel, the city centre will see a combination of LiDARs, computing units, communication units, and using data analytic techniques to track pedestrian movement patterns and density. The data collected will be crucial for improving pedestrian safety, managing crowd flow, and informing city planning initiatives. The back-end data centre can obtain real-time processing results at the Edge and visualize them to the public. As the pilot progresses, the data gathered will be evaluated and used to adjust traffic patterns, modify transport routes, and potentially redesign city layouts to better accommodate pedestrian and vehicular flow.

This pilot scenario, combining technological advancement and data-driven decision-making, is the first step towards transforming Göttingen into a truly smart city, enhancing the quality of life for its residents and visitors alike.

Edge nodes co-located with the cameras represent Kubernetes (K8s) worker nodes; the control plane is expected to reside in the Cloud. Hence, in the context of this use-case, CODECO shall be used to orchestrate (reallocate) resources across Edge-based environments, to assist in the degree of control decentralisation.

2.4.1.2 Source

HE-CODECO project

2.4.1.3 Roles and Actors

Actors:

GOV and municipalities can monitor traffic and get analytics on patterns (e.g., traffic, volumes of cars, bikes, etc.).

Network Infrastructure provider offers network connection.

Cloud/Edge infrastructure provider offers computation and storing resource for analytics.

User (developer, subscriber, e.g., Citizen) can get information about traffic.

Roles:

Transportation planning and management: The government and municipalities will have access to real-time traffic analytics, including abnormal alarms, as well as long-term traffic pattern analytics. By automatically collecting and analysing data on traffic behaviour, they can gain insights that can inform decision-making and improve city planning, enhancing traffic safety and efficiency.

Property development and investment: The infrastructure provider and deployer need to carefully consider the selection, deployment, and maintenance of the system, considering the feasibility and costs associated with each step. Once the system is deployed, the infrastructure provider conducts testing and validation to ensure that it is working effectively and providing the desired results.

Advertising and marketing: The citizen will have access to real-time traffic information through downloadable apps or public boards, enabling them to plan their travel and make informed decisions about their routes.

A summary of the business impact for the different proposed user journeys is as follows:

User Journey #1 Transportation planning and management: The real-time traffic data collected by the road usage monitoring system can be used to optimize transportation planning and management in the city. This can benefit businesses that rely on efficient transportation and logistics, such as delivery services, public transportation companies, and trucking companies.

User Journey #2 Property development and investment: The pedestrian distribution monitoring part of the system can provide valuable insights into foot traffic patterns in different areas of the city.

This information can be used to make data-driven decisions about property development and investment. Real estate companies and property developers can use this data to identify high-traffic areas and make informed decisions about where to invest in new properties or develop existing ones.

User journey #3 Advertising and marketing: The data collected by the system can also be used for advertising and marketing purposes. For example, businesses can use the data to identify high-traffic areas and strategically place their ads in these locations. This can help businesses reach their target audience more effectively and potentially increase sales.

2.4.1.4 Pre-conditions

Specific equipment and Edge nodes placed across the city in specific locations and are dimensioned with specific constrains in terms of computational and networking resources.

Application workload and respective datasets need to be uploaded, or adequate traffic needs to be generated e.g., by the installed cameras.

2.4.1.5 Triggers

Specific nodes in the infrastructure go down, or new nodes enter the living system – critical changes to the infrastructure.

2.4.1.6 Normal Flow

Figure 28 provides a high-level perspective on the proposed system architecture considering one cluster deployment, detailing both non-CODECO and CODECO components. The K8s control plane with the respective CODECO nodes is represented to run in the Cloud but may also be deployed at the near Edge. The worker nodes are deployed in Edge nodes co-located with cameras across the city.
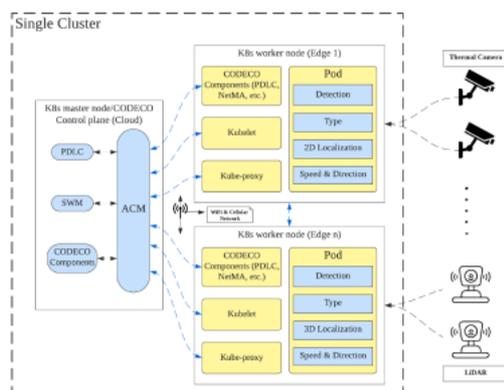
Figure 28: CODECO P1 system architecture.

As illustrated, there are CODECO components and non-codeco components.

Non-CODECO components deliver highly accurate and timely traffic indicators such as vehicle counting, traffic density, and traffic flow by using thermal camera sensors. With this data, cities can have an overview of the number of vehicles entering/leaving Göttingen, monitor traffic patterns, and use the information to reduce congestion. This information can also be used to identify potential bottlenecks or problem areas in the transportation system, allowing officials to proactively address issues before they become major problems. In addition, this information could be used to optimize existing infrastructure to better meet the needs of residents and visitors. These components will also deliver highly accurate and timely pedestrian distribution indicators such as pedestrian counting, pedestrian density, and pedestrian distribution heat map by using the LiDAR sensor. It can be a valuable tool for city planning. On the one hand, it can help the city to have an overview of the number of pedestrians in some key areas in Göttingen city centre. On the other hand, it can help to understand pedestrian traffic patterns: Heat maps can provide the city with a visual representation of where pedestrians tend to congregate, move, and dwell in public spaces. By analysing these patterns, the city can identify high-traffic areas and prioritize them for improvements, such as new crosswalks, street furniture, or public amenities.

CODECO Components shall be placed both at the Cloud (K8s master node) and Edges (K8s worker nodes). For each location chosen for implementation, we will outfit it with an Edge device equipped with an array of sensors such as thermal cameras or LiDAR systems. In this use-case. An Edge node will be the same as a K8s worker node. This arrangement ensures the autonomous computational and data processing capabilities of each Edge device, all while remaining interconnected within the larger network. The data collected by these devices, including video and point Cloud data, must be pre-processed, and stored locally. This local storage of data is a crucial step that allows for future reference, audits, and given the sensitivity of the data to outflows from the specified area.

2.4.1.7 Alternative Flow

N/A

2.4.1.8 Post-conditions

Once the risks have been mitigated or circumvented, the system continues in operation. The user gets information via a dashboard; but the change in the infrastructure is agnostic to the user.

2.4.1.9 High Level Illustration

See **Figure 28**.

2.4.1.10 Potential Requirements

2.4.1.10.1 Deployment KPIs

Minimize bandwidth costs: we want to consider the bandwidth usage of each compute node and factor that into our decision-making process when deciding where to place workloads. We may also want to consider the bandwidth costs associated with network connections, such as those incurred when data is transferred between nodes. Overall, our goal would be to choose a deployment strategy that minimizes our bandwidth costs while still meeting our performance and reliability requirements.

Bandwidth cost: < 10G/month/camera

Overall detection and counting accuracy: >= 80%

System latency: <= 40ms

Rate of occurrence of failure (ROCOF): <= 5%

2.4.1.10.2 Non-functional requirements

Privacy protection in video collection and transmission; GDPR Compliance.

2.4.2 CODECO P2: Vehicular Digital Twin for Safe Urban Mobility

Contact: I2CAT, Jordi Marias (jordi.marias@i2cat.net)

2.4.2.1 Description

P2 makes use of the CODECO framework to support a Vehicular Digital Twin aimed to improve the safety of *Vulnerable Road Users (VRU)* in Urban Environments. Any mobility oriented Digital Twin requires the extensive deployment of ultra-reliable low latency services around the area it supports. Starting from the V2X communication capabilities to Computer Vision (CV) detectors capable of tracking all the moving parts within the mobility environment.

For this reason, the current use case relies on *V2X Roadside Units (RSUs)* and cameras to gather all the necessary information to track vehicles and pedestrians and then feed it to the vehicular Digital Twin, which will detect dangerous situations or behaviours and alert them.

The deployment and scalability of this service has challenges around the infrastructure side, where the information should be processed as close as possible to the V2X nodes and ensure low latency communications. This, in turn, translates to always having a fresh track of all the moving parts.

The pilot scenario focuses on the mobility environment of the interior and adjacent street of a UPC campus known as "Campus Nord" in Barcelona, which is located next to the I2CATs offices. This environment offers an interesting balance with walkable pedestrian zones that include bike lanes, as well as car lanes on the adjacent street. It encompasses a mix of various transportation modes, with *Vulnerable Road Users (VRUs)* playing a central role. However, VRUs can find themselves in dangerous situations when sharing spaces with cars. Which addresses the UC being presented.

This scenario presents an ideal testing ground due to its size, allowing for the examination of multiple areas and their respective control measures. Additionally, it provides a diverse representation of all transportation modes commonly found in urban environments. Consequently, this scenario offers the perfect setting to assess and address the challenges associated with different modes of transportation, ensuring the safety and efficiency of urban transportation networks.

2.4.2.2 Source

HE-CODECO project

2.4.2.3 Roles and Actors

Actors:

VRUs. A VRU, refers to vehicles or transport users who lack physical protection while navigating roadways, making them more susceptible to harm. They require heightened caution from both them and drivers to ensure their safety on the streets. Their role is to move around and expect to be advised when engaged in dangerous situations. Their goal is to prove that they can be safer thanks to the infrastructure services and the notification through V2X.

Pedestrians: They will be equipped with a smartphone and tracked by the camera. Are the most vulnerable ones.

Light Mobility Vehicles (bike or electric scooter): Equipped with an OBU or smartphone and tracked by a camera. They tend to get in riskier situations due to its speed and fragility.

Cars: Equipped with an OBU. Can find themselves in dangerous situations when sharing mobility spaces with lighter, more vulnerable, modes of transportation. Their role is to move around the car lanes and engage in risky behaviours with VRUs and to proof that the dangerous situations are avoided when detected by the system and notified.

Cameras: The cameras, which will be connected to the corresponding service. Will spot VRUs that are not connected. And notify them of their position and track to the infrastructure. Their role is to detect the position and trajectory of all the moving entities that are not connected.

To then pass this information to the infrastructure. Their goal is to make sure the use case works even when there are some vehicles or pedestrians not connected.

V2X RSUs. They will be the point of contact from the vehicles and UC users and the infrastructure. There will be some of them to ensure full coverage of the use case scenario. They will need to ensure minimum latency to the infrastructure. Their role is to serve as point of contact between the infrastructure and the V2X environment. Their goal is to make sure the messages emitted by the OBUs, or the infrastructure are mutually received.

Roles:

Governmental entities, policy makers, will have the availability to see a real time low-latency digital twin of the moving entities around the designated urban public space. Check dangerous behaviour and immediate notifications when any incident happens (along with the place where the incident happened and how it happened).

ICT (industry, mobile communications, and Cloud providers) will use CODECO to achieve reliability and low latency.

The **academic and researcher stakeholders** will be capable of ensuring through the proper research the viability of a system of Vulnerable Road User Awareness. And provide solid information that reduces/eliminates incidents with that kind of users.

The **developers** (Edge, IoT) and especially the **early adopters** will run around the campus as VRU. And will receive an audio-visual notification. On the other side other early adopters will be using conventional cars that will also notify collisions.

A summary of the business impact expected is as follows:

One potential business case for this technology is in the public sector. Municipalities and other government agencies responsible for mobility could use this technology to monitor and reduce the number of incidents involving vulnerable road users and other vehicles. By providing real-time data on the movements of these users, the application could help city planners design safer roads and intersections.

Another potential business case is in the insurance industry. Insurance companies could use the data collected by the application to improve their revenue by accurately assessing risk and reducing the number of claims. The positional and camera data provided by the application could also help insurance companies more accurately determine fault in the event of an accident.

The automotive industry is also a potential market for this technology. Car manufacturers are always looking for new features to include in their vehicles, and a V2X Vulnerable Road User awareness application could provide an extra layer of safety for drivers. This technology could be particularly useful for vulnerable vehicles such as bikes, electric scooters, and motorcycles.

2.4.2.4 Pre-conditions

Object detection (cars, pedestrians).

2.4.2.5 Triggers

Movement changes, time-based.

2.4.2.6 Normal Flow

The heterogenous sensor-V2X related modules of the present use case interact with each other following the structure shown on **Figure 29**.
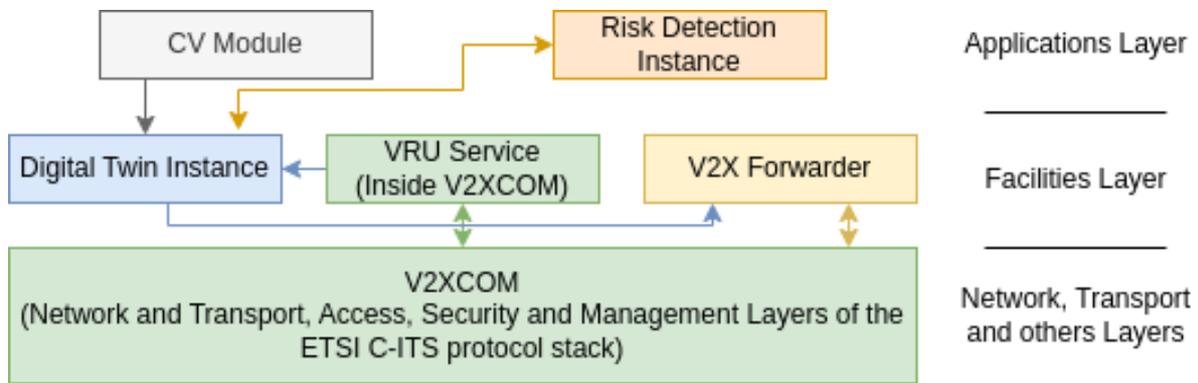
Figure 29: P2 UML diagram

In addition to the functionalities described in the previous section, it is essential to outline the various information flows between the modules:

1. **V2XCOM to V2X Forwarder**: All packets, regardless of the Facilities standard they belong to, are transmitted to the V2X Forwarder. Each packet is accompanied by information specifying the Radio Access Technology (RAT) it originates from or the unicast IP in the case of conventional mobile communications. The V2X Forwarder then determines whether to forward the message to other RATs and unicast Ips (using the V2XCOM module) or ignore it altogether.

2. **Digital Twin to V2X Forwarder**: Certain packets received from the V2XCOM module may be ignored by the V2X Forwarder. This occurs because the V2X Forwarder awaits queries from the Digital Twin to aggregate multiple messages and achieve the same effect as directly forwarding the packets.

3. **VRU Service (V2XCOM) to Digital Twin**: Whenever a new Vulnerable Road User (VRU) Awareness Message (VAM) is received, it undergoes processing, and its information is directly stored in the Digital Twin. Digital Twin serves as a repository for VRU-related data.

4. **Digital Twin to Risk Detection Instance**: The Risk Detection Instance continually receives updates on the positions and trajectories of all moving entities from the Digital Twin. With each update, the Risk Detection Instance assesses the potential for dangerous situations. If a hazardous situation is identified, the Risk Detection Instance notifies relevant parties.

5. **CV Module to Digital Twin**: The CV Module continuously processes images captured by associated cameras and extracts positional information from the detected nodes. This positional information is promptly stored in the Digital Twin, allowing for comprehensive data integration.

These information flows ensure efficient communication and coordination between the modules, enabling the exchange of crucial data for the proper functioning of the system.

2.4.2.7 Alternative Flow

N/A

2.4.2.8 Post-conditions

Once the risks have been mitigated or circumvented, the system continues in operation. The user gets information via a dashboard; but the change in the infrastructure is agnostic to the user.

2.4.2.9 High Level Illustration

See **Figure 29**.

2.4.2.10 Potential Requirements

2.4.2.10.1 Deployment KPIs:

Latency: Average latency (V2V, V2P or V2I) < 50ms.

Age of Information (AoI): Average AoI < 70ms; Average Peak AoI < 200ms

Penalty Age of Information: The penalty function is the L2-Norm between the estimated trajectory and the real one.

Average Penalty Age of Information: < 20.

Average Peak Penalty Age of Information: < 50.

Processing time: < 5 ms.

Neighbourhood Awareness Ratio: must be 100% for distances < 100 meters. There is a tolerance of missing a neighbour for distances from 100 to 300 meters. (Nar above 80%).

2.4.3   CODECO P3: MDS across Decentralised Edge-Cloud

Contact: Telefonica, Luis Contreras Murillo (luismiguel.contrerasmurillo@telefonica.com)

2.4.3.1 Description

P3 focuses on the smart and efficient distribution of media content (e.g., video streaming, gaming, Augmented Reality/Extended reality (AR&ER) across a multi-domain, multi-cluster Edge-Cloud. The use-case therefore leverages on a combined optimization of both connectivity (from the underlying transport network) and computational resources (supporting the MDS streamers and distribution logic).

P3 promotes a tighter computational/networking integration and optimizes the overall resource usage while reaching a good level of *Quality of Experience (QoE)*. To reach this, the use-case focuses on an interaction between a *Media Delivery System (MDS),* via CODECO, and the CODECO component NetMA which shall rely on a decentralized concept of the IETF ALTO protocol[4] to expose capabilities (e.g., topological information together with associated metrics, available resources, or functions) that promote joint adaptation.

The key aspects to be demonstrated concern:

- demonstrating the CODECO informed orchestration of virtualized delivery points with the purpose of selecting the most appropriate Edge facility- according to specific constraints in both the Edge-compute (CPU, RAM, or storage) and the network sides (i.e., latency, bandwidth),

- obtaining a real-time, updated view of the network status, for instance due to optimizing the delivery, for triggering on-demand instantiation of Edge delivery points.

2.4.3.2 Source

HE-CODECO project

2.4.3.3 Roles and Actors

Actors:

MDS platform owner.

Network Infrastructure owner.

Cloud/Edge infrastructure owner.

---

[4] https://datatracker.ietf.org/wg/alto/about/

MDS content owner.

MDS subscriber.

Roles:

The primary actor is the application-service provider (and network capability client). There are also two secondary actors: the application-service client, and the network operator.

A summary of the business impact is as follows:

To fulfil the ever-changing demands of their clients, companies in the telecommunications sector are continually working to improve their services. Customer turnover is one of the most important issues that service companies must deal with. The revenue and profitability of a firm can be significantly impacted by high customer turnover rates. Service providers are making investments in technology that can raise the quality of the customer experience (QoE) to deal with this problem.

One way to reduce costs without a great investment is optimizing the use of the resources available. To archive this, this use case uses exposure capabilities to allow the network client to select the best path according to the nodes and path characteristics and the internal client information. By this, the telecom operator increases service satisfaction improving QoE. This solution allows better resource use, reducing delivery expenses and optimizing the capabilities available in both network and Edge-Cloud.

2.4.3.4 Pre-conditions

Application workload and respective datasets need to be provided.

Pre-authorization of users needs to be in place.

2.4.3.5 Triggers

Specific nodes in the infrastructure go down, or new nodes enter the living system – critical changes to the infrastructure.

Workload migration is triggered by CODECO (SWM component).

2.4.3.6 Normal Flow

For the CODECO MDS use case, an initial system architecture is defined, which is subject of refinement along the implementation of the use case and the development of the different CODECO components.

For each of the service situations, i.e., single-cluster and multi-cluster, two different system behaviours are exercised.

In the single cluster, understood as baseline case, the MDS will interact with the CODECO components with the objective of optimizing the content delivery considering network metrics in addition to the MDS view of the delivery resources. Such interaction between MDS and CODECO will imply the interaction with the different components (NetMA, SWM, MDM, etc) for multiple purposes. In the following figures, only the NetMA component is highlighted for the only purpose of remarking the interaction with this component for retrieving network information (by leveraging on ALTO).

The operation of P3 in a single cluster is represented in **Figure 30**. In the single-cluster case, once the end-users request a service to the MDS (step 1), the MDS logic interacts with the CODECO components for retrieving information about the network metrics applicable to the different existing delivery points (step 2). Once this information is retrieved, the MDS analyses the combined compute and network information, including the availability of contents, and takes a decision on what is the more convenient delivery point to serve the end-users (step 3). With that, the end-users access the selected delivery point.

On the other hand, the multi-cluster case is exercise for the optimal instantiation of delivery points across the Edge-Cloud continuum motivated by the presence of end-users in each area. As before, the interaction between MDS and CODECO components implies the interplay with different components, even though only NetMA is highlighted for simplicity.
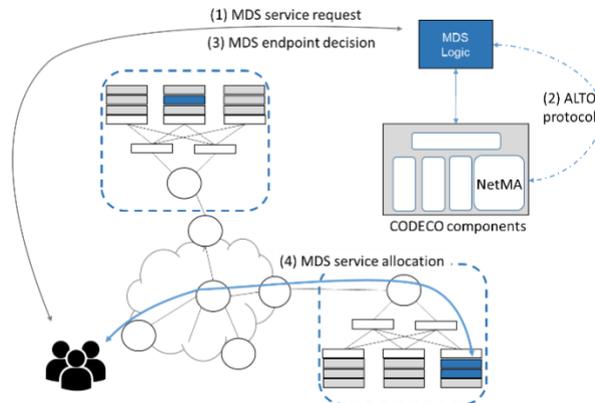

Figure 30: P3 single cluster representation

In the multi-cluster case, the MDS identifies a (potential) end-user base that can drive an optimization of the delivery system (step 1). To identify the more convenient node where deploy a new delivery endpoint in the Edge-Cloud continuum, the MDS logic interacts with the CODECO components to retrieve potential sites where deploy the endpoint collecting both network and compute metrics (step 2). After processing all that data, the MDS logic will decide where to instantiate the new delivery point, optimized from that perspective. The new endpoint will be interconnected with the Origin MDS node as well with the rest of the MDS footprint by means of L2S-M overlay for feeding contents, etc. (step 3). Once the new endpoint is available, the requests coming from the end-users will be served as described in the single-cluster case.

Note that the motivations for the placement of new delivery endpoints can be triggered by other processes and situations, e.g., as for a workload migration triggered by SWM.

As said, the use case will be refined along the progress of the project, so different situations could be documented as part of the single- and multi-cluster cases.
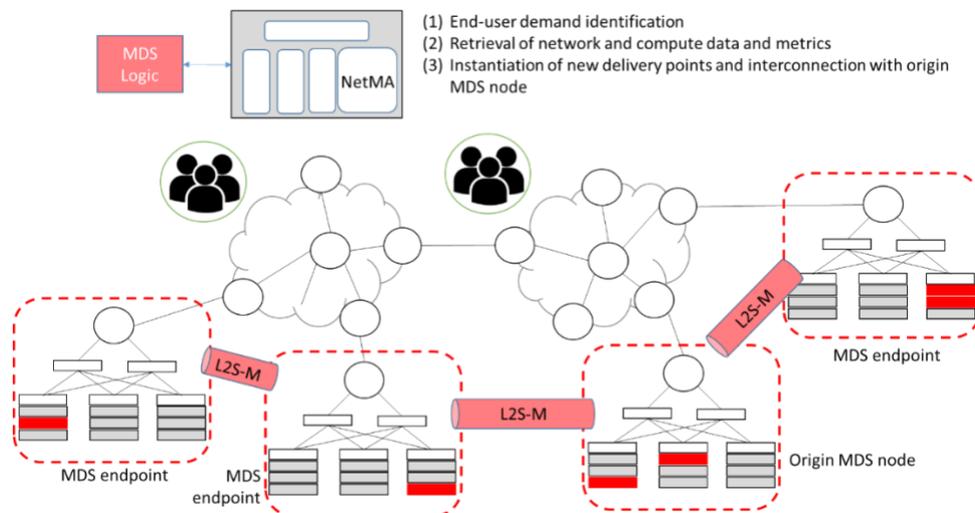
Figure 31: P3 multi-cluster architecture representation

2.4.3.7 Alternative Flow

N/A

2.4.3.8 Post-conditions

After the instantiation of new delivery points to serve end-users, content delivery is optimized and proceeds as indicated in the service.

2.4.3.9 High Level Illustration

See **Figure 30** and **Figure 31**.

2.4.3.10 Potential Requirements

2.4.3.10.1 Deployment KPIs:

Efficiency level achieved in the usage of both computer and network resources.

Reaction time reduction in terms of adaptation execution.

2.4.3.10.2 Non-functional requirements:

Endpoint Authentication: Even the information is not critical, it is needed an endpoint authentication to avoid data poisoning.

Interoperability: System should be multi-vendor and do not have dependencies with the hardware/software used to deploy the UC.

Maintenance: System should be easy to maintain and update, having clear documentation about its parts, how it works and how to update it.

Reliability: System should be able to detect a fall and recover from it. Also, should have a failure-resistant deployment.

2.4.4   CODECO P4: Demand-side Management in Decentralized Grids

Contact: UPM, David Jimenez (djb@gatv.ssr.upm.es)

2.4.4.1 Description

The proposed use case for the distributed energy management system focuses on implementing an active demand response decentralized management system for building decarbonization. It aims to optimize energy usage, improve sustainability, and enhance the resilience of buildings by integrating renewable energy sources and enabling intelligent demand response actions.

The use case also emphasizes the joint orchestration of computational and networking resources to ensure efficient coordination and management of energy-consuming devices and the networking infrastructure within buildings. It focuses on achieving a holistic view of data in the IoT-Edge-Cloud continuum, enabling comprehensive monitoring, analysis, and replication of energy-related data.

CODECO framework leverages the power of K8s to build a decentralized energy management system. By integrating worker nodes (which in this use-case have a correspondence with Edge nodes) and employing the CODECO's developed tools like ACM, PDLC, and modular functionalities, P4 aims at achieving efficient resource utilization, scalability, resilience, and adaptability in energy management operations integrating the energy-related IoT systems and the computing needs.

2.4.4.2 Source

HE-CODECO project

2.4.4.3 Roles and Actors

Actors:

Prosumers

Energy Communities

Aggregators

DSO

Cloud/infrastructure providers

Roles:

Energy service provider: UPM aims to be climate neutral by 2030. To this end, it seeks synergies between generation and energy consumption between its different campuses and its buildings. It leverages Edge computing capacity to make its associated generation and consumption forecasts, and scheduling and optimization models. From them, it decides how to group the different energy assets to achieve the best performance in terms of decarbonisation. This calculation is executed in the Cloud and associated to the physical environment (cluster) defined, where IoT data collected on the real-time operation to take the necessary corrective measures on planning.

Energy community: UPM acting as an energy community wants to offer energy aggregation (offering network flexibility and balance services to the system operator and distributors) and charging services for electric vehicles.

A summary of the business impact is as follows:

The business case for a distributed energy management system lies in its ability to optimize energy usage, enhance grid resilience, and enable the integration of renewable energy sources. By decentralizing energy management, such a system allows for efficient utilization of distributed energy resources, including solar panels, wind turbines, and energy storage devices.

With increasing concerns about climate change and the need to transition to a low-carbon economy, the market demand for distributed energy management systems is growing. This demand is driven by factors such as government regulations promoting renewable energy adoption, rising energy costs, and the desire for energy independence and resilience.

The market analysis reveals a significant potential for growth in the distributed energy management system market. The system caters to various sectors including residential, commercial, and industrial, where energy consumers seek ways to reduce costs, improve sustainability, and contribute to environmental goals. Additionally, the integration of advanced technologies, such as IoT, AI, and blockchain, into these systems further enhances their capabilities and market attractiveness.

Key market players in the distributed energy management sector include energy service companies, technology providers, utilities, and energy aggregators. These players offer a range of solutions, including energy monitoring and control platforms, demand response management systems, and virtual power plant solutions.

The impact of distributed energy management systems is multifaceted. They enable consumers to reduce their energy bills through optimized energy usage and by leveraging locally generated renewable energy.

Additionally, these systems contribute to grid stability and resilience by balancing energy supply and demand in real-time, reducing the strain on centralized power infrastructure. Moreover, distributed energy management systems foster the integration of renewable energy sources, accelerating the decarbonization of the energy sector. They facilitate the transition from a traditional centralized energy model to a decentralized and democratized energy system.

Overall, the high-level market analysis reveals a growing demand for distributed energy management systems driven by energy cost savings, sustainability goals, and the need for resilient energy infrastructure. The market presents opportunities for innovative solutions and collaboration among stakeholders to transform the energy landscape towards a more sustainable and efficient future.

One way to reduce costs without a great investment is optimizing the use of the resources available. To archive this, this use case uses exposure capabilities to allow the network client to select the best path according to the nodes and path characteristics and the internal client information. By this, the telecom operator increases service satisfaction improving QoE. This solution allows better resource use, reducing delivery expenses and optimizing the capabilities available in both network and Edge-Cloud.

2.4.4.4 Pre-conditions

Application workload and respective datasets need to be provided.

Pre-authorization of users' needs to be in place.

2.4.4.5 Triggers

CODECO alerts on abnormal pattern detection

On-demand requests for optimization.

2.4.4.6 Normal Flow

The system architecture represented in **Figure 32** is still in a very initial stage of development. Nonetheless it shall rely on CODECO, leveraging a combination of Edge computing, Cloud computing, and K8s for efficient resource allocation, data processing, and decision-making. The architecture consists of the following components:

Edge Devices: These are devices located at various energy generation and consumption points, equipped with sensors, and connected to the local energy infrastructure. Examples include smart meters, renewable energy sources (e.g., solar panels, wind turbines), and energy storage systems. Each Edge device acts as a worker node within the K8s system, capable of performing computations and data processing independently.

Master Node: The master node serves as the central control plane within the K8s system. It manages and orchestrates the distributed resources and tasks across the Edge devices. The master node is responsible for coordinating energy generation, consumption, and optimization algorithms, as well as collecting and analysing data from the Edge devices.

Energy UC Cluster: The Energy UC cluster consists of the master node and multiple worker nodes (Edge devices). It provides a scalable and resilient infrastructure for managing the decentralized energy system. The K8s cluster handles workload scheduling, resource allocation, and load balancing to optimize energy management tasks.
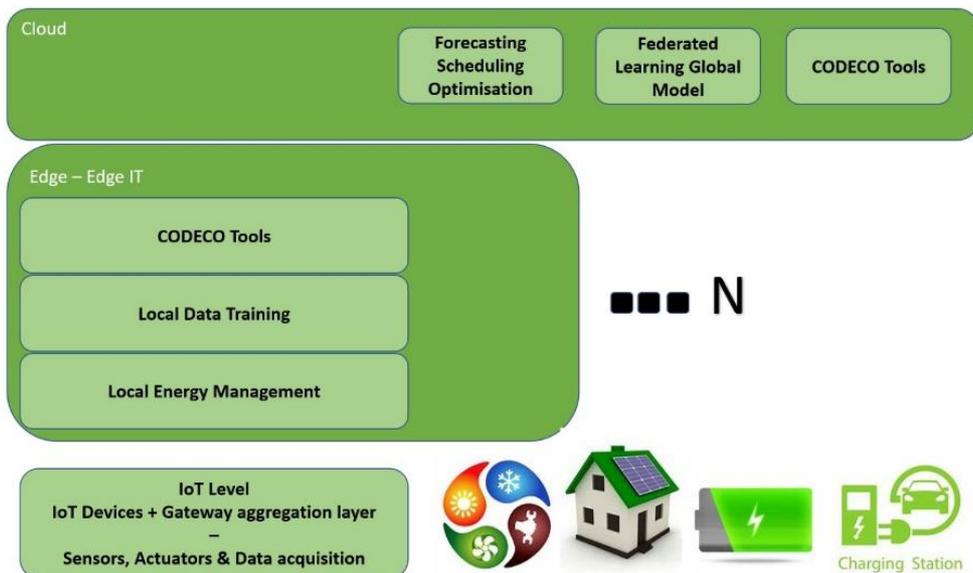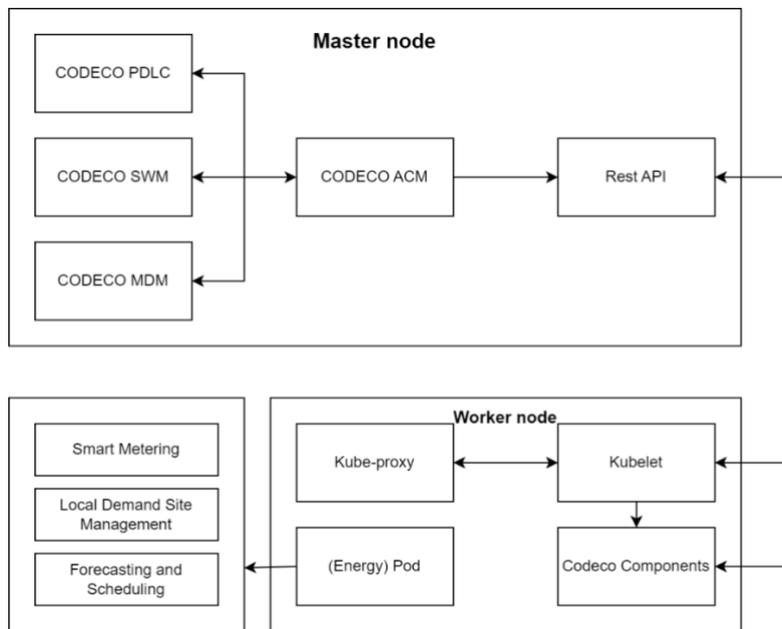

Figure 32: P4 system architecture

The initial proposed workflow for P4 is illustrated in **Figure 33**, for a single cluster deployment.

Figure 33: P4 UML use-case diagram

This UML diagram illustrates the flow of data and control in a decentralized energy demand side management system. The sensors collect energy-related data, which is transmitted to the Edge device for local management execution. The local manager handles real-time decision-making and control tasks. For more complex optimization functions, some computing is offloaded to the Cloud, where advanced algorithms or machine learning models can be applied to optimize energy consumption, load balancing, or other energy-related tasks. Internally in each of the pods, there are three specific use-case components:

Smart Metering.

Local Demand Site Management.

Forecasting and Scheduling.

2.4.4.7 Alternative Flow

N/A

2.4.4.8 Post-conditions

Optimized energy consumption.

2.4.4.9 High Level Illustration

See **Figure 32** and **Figure 33**.

2.4.4.10 Potential Requirements

2.4.4.10.1 Deployment KPIs:

Number of buildings involved. (>3)

Energy assets integrated. (>100)

Amount of kilowatt and kilowatt hours of the energy community. (30% UPM consumption)

Amount of energy saved (not bought from the grid). 10-20%

Number of energy clusters created and modified per day. (20)

CO_2 emissions reduction. (>10%)

2.4.4.10.2 Non-functional requirements:

Trusted Execution Environments (TEE)

2.4.5   CODECO P5: Wireless AGV Control in Flexible Factories

Contact: Rute C. Sofia (sofia@fortiss.org)

2.4.5.1 Description

There is today an increasing need to consider Automated Mobile Robots (AMRs), of which one example are Automated Guided Vehicles (AGVs) in manufacturing environments, to support the heterogeneous and growing demand of material handling and logistics in flexible factory environments.

While current AGV fleets are based on pre-defined task assignment and pre-defined paths, there is an urgent need to provide a more flexible control to support fleets with a larger number of AGVs, and to support an increasing number of tasks/goods to be transported. By reaching a higher level of autonomy, it is possible to increase overall efficiency while reducing operational costs. The integration of wireless technologies to support the control of AGVs, e.g., 5G, Wi-Fi 6/7, becomes highly relevant and shall be explored by CODECO. However, relying on wireless implies also that the control of AGV systems is prone to interference and intermittent connectivity, thus requiring a higher degree of adaptation which CODECO is expected to provide.

Hence, in addition to the wireless connectivity aspects concerning interference mitigation, synchronization, this use-case shall also demonstrate the CODECO capability to proactively adapt the overall network energy consumption and to mitigate interference and failures.

In this context, the use-case shall explore AGVs handling goods within a warehouse, being subject to remote control and requiring real-time support. The use-case expects to be developed in three phases:

Phase 1, single cluster, static control plane.

Phase 2, single cluster (multi-master), mobile control plane.

Phase 3, federated clusters.

AGVs shall therefore expected to carry different micro-services (dockerized) for a single cluster. In this case AGVs shall correspond to Kubernetes worker nodes, while the control plane shall reside on a static node. The AGV micro-services shall be managed via CODECO, being the CODECO components placed across the control and data plane of K8s. On a second phase, the control plane shall also be deployed on a mobile node.

CODECO shall explore distributed ML approaches considering computation as close as possible to data sources; networking features (e.g., available bandwidth); energy awareness, to assist in a higher degree of autonomy. The CODECO framework will be installed across the fortiss IIoT Lab (expected to reach 10 nodes, mobile and embedded).

The experimental environment will be developed based on realistic scenarios, derived from consultation with manufacturing partners. A final demonstration involving multi-cluster domains will be provided together with an external manufacturing partner of FOR.

2.4.5.2 Source

HE-CODECO project

2.4.5.3 Roles and Actors

Actors:

AGVs (far Edge nodes) – mobile robots with different sensors (e.g., cameras, environmental sensors)

User: user DEV, developer willing to deploy the CODECO AGV App; user Operator, human operators, and respective terminals, remotely assisting AGVs.

AGV fleet Controller node.

Roles:

Deployment of micro-services in an AGV fleet: a user (DEV, developer) wants to deploy a new CODECO offered application in UC5 across an AGV fleet and wants also to manage its application workload with K8s/CODECO.

The user (DEV) shall be able to observe the existing cluster via a CODECO dashboard (9), and be able to make initial adjustments, if required (9).

AGV Fleet control – Resilient infrastructure: This user journey relates with the runtime management of the CODECO AGV Apps. The aim is to assist AGVs in autonomous navigation on indoor, blocked spaces. Key challenges concern energy optimization and support of intermittent connectivity. ICT stakeholders relevant for this use-case are mobile communications, Edge-Cloud providers.

A summary of the business impact is as follows:

The P5 value proposition (VP) canvas is provided in **Figure 34**. The application of CODECO to the context of AGV fleet decentralized control has as customer segments the CODECO target groups DEV (developers), ICT (large industry and SMEs) and AR (Academia and Research). The targeted vertical domains are Manufacturing and Logistics, which correspond to domains where there is an increasing growth in the need of automation and cognitive processes to improve the overall operations in critical environments. With the integration of Industrial IoT and ML, these sectors are experiencing a major change towards decentralisation, as observable in the concept of Manufacturing as a Service (MaaS).

Target customer	VP 1	VP2
Customer perspective	DEV: segment wants zeroconf deployment, low skill investment	ICT, AR, SMEs: zeroconf, scalability, and low cost fleet management, with energy efficiency
Competing alternative	Proprietary solution providing 99.999% reliability High investment in training	Proprietary, customized solution often tailoring 1 cluster. Federated clusters (e.g., remote locations) requires high investment (CapEx and OpEx)
Differentiators		
Performance	**Scalability**, capability of CODECO to cope with an increasing number of application deployments across variable fleet sizes (large, heterogeneous) and towards mobile devices. **Resilience and availability**, capability of CODECO to support five nines system availability in the verge of network interference and intermittent connectivity.	**Privacy,** capability of CODECO to manage a varying infrastructure across a single or different locations (multi-domain environments). **Resilience and availability**, capability of CODECO to support five nines system availability in the verge of network interference and intermittent connectivity.
KPIs	• Time to completion of a task (latency): reduction in 20% due to decentralized control. • 10% of reduction in the number of collisions. • 10% improvement of energy efficiency of the network (overall involved nodes) and eventual network lifetime.	10% improvement of energy efficiency of the network (overall involved nodes) and eventual network lifetime. 10% improvement of total setup times

Figure 34: P5 value-proposition canvas

The proposed solution in this use-case consists of CODECO and of a set of AGV fleet Apps to assist the deployment of the use-case, and to play with CODECO components. The key innovation aspects in UC5 relate with the use of context-awareness and behaviour estimation to provide a higher degree of flexibility to the overall system, thus allowing control of AGVs to be handled in a decentralized way, expected to bring benefits in large-scale environments.

In terms of performance, the application of CODECO in UC5 is expected to improve scalability, resilience, and availability in comparison to K8s, adding also novel support in mobile environments.

2.4.5.4 Pre-conditions

Active AGV fleet interconnected via wireless.

2.4.5.5 Triggers

Node (AGV) with intermittent connectivity.

Node (AGV) goes down due to battery or another failure.

Node (AGV) not capable of supporting an assigned task (e.g., lack of memory, energy, etc).
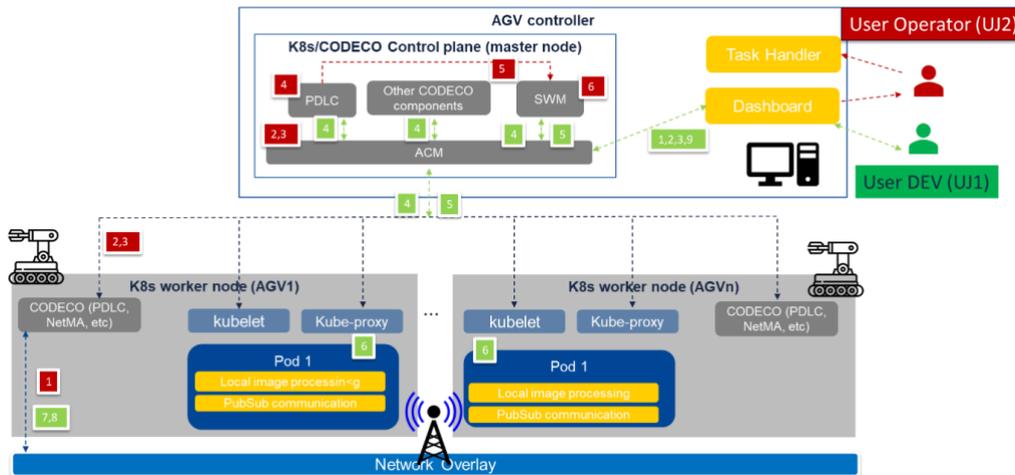
Network with low link quality.

2.4.5.6 Normal Flow

Figure 35: P5 system architecture, one cluster

Deployment of micro-services in an AGV fleet

Rf. To Figure 15. The user wants to deploy a new CODECO offered application in UC5 across an AGV fleet and wants also to manage its application workload with K8s/CODECO. For this purpose, the user starts by accessing the CODECO ACM UI (1) via the available dashboard (AGV controller). The dashboard shall interact with the ACM UI, via a specific customization for the use-case. Hence, via the dashboard the user DEV shall be able to select a pre-defined set of micro-services deployed for the use-case (2). For the initial CODECO AGV App, the UC shall provide a basic set of micro-services available, such as the ones illustrated in Figure 2. For instance, PubSub approach such as MQTT Sparkplug or NDN; micro-service for task handling; micro-service for object detection). Some of these services will be mandatory; some will be optional.

Then, the user is also requested to enter a set of requirements (3), e.g., key requirements such as latency; size of the fleet (how many AGVs to consider); type of communication (e.g., 5G, Wi-Fi); channel aspects, etc. These parameters serve the purpose of creating the so-called CODECO Application model (YAML file(s)), which is key to adequately schedule resources to be used.

Once completed, ACM stores this information (ApplicationModel, CRD format, (4)), making it accessible via the usual K8s methods to other components of CODECO, in alignment with the CODECO CRs/CRDs. SWM starts the initial placement (5). The deployment of the AGV services (ApplicationGroup) is started, being all deployment developed in a single cluster (1 Pod per worker node; 1 AGV corresponding to a worker node) set up by default with all involved nodes that are within range at an instant in time, and that may appear later in the radio range of the controller (6).

For the case of an AGV fleet based on multiple remote locations (phase 2), then ACM shall activate the procedures for federated clusters, instead of deployment on a single cluster. Further development aspects shall be considered during the development of CODECO features for federated clusters (M18-M36).

CODECO shall handle in addition the required network path handling, by taking into consideration aspects such as interference mitigation, channel properties, etc. This shall be handled via the information collected via NetMA for the wireless interconnections across AGVs (7). If required, routes shall be set to optimize the overall communication (8).

The user (DEV) shall be able to observe the existing cluster via a CODECO dashboard (9), and be able to make initial adjustments, if required (9).

AGV Fleet control – Resilient infrastructure

This user journey relates with the runtime management of the CODECO AGV Apps. The aim is to assist AGVs in autonomous navigation on indoor, blocked spaces. Key challenges concern energy optimization and support of intermittent connectivity. ICT stakeholders relevant for this use-case are mobile communications, Edge-Cloud providers.

For this deployment we will investigate existing proposals for AGV communication, e.g., derived from VDMA guidelines and shall consider both a single cluster and a multi-cluster deployment.

On a first phase, we shall consider a centralized approach where the central controller has a global perspective on the overall K8s infrastructure (data, compute, network, (1)) which is regularly updated based on data collected via different CODECO components and managed via the CODECO CRs/operators (2, 3). The CODECO PDLC performs, for this specific scenario, an analysis of robustness of the overall graph, and of the existing links in terms of energy consumption across a pair of nodes, as well as in terms of channel conditions, RTT, between two nodes (4). It can propose an adaptation of the overall communication infrastructure derived from functional and non-functional network requirements to the SWM scheduler (5) which shall then decide on whether to adapt the overall infrastructure (6). Additional re-scheduling supported by CODECO shall take into consideration aspects such as energy consumption. If an AGV is expected to run out of battery in x seconds, then its micro-services shall be passed (replicated or offloaded) to another suitable AGV, automatically selected by CODECO based on the Application model requirements provided by the user.

On a second phase, we shall consider a decentralized approach, where each AGV shall be responsible to transmit its own perspective of the K8s infrastructure at an instant in time to other AGVs. The infrastructure data (data observability, computation, network) is regularly updated by different CODECO operators (2) to the CODECO control plane, which now shall consist of a multi-master cluster. The selection of 3 master nodes per cluster, to ensure resilience, shall be done based on NetMA input (1) to ensure a stronger resilience to failures.

2.4.5.7 Alternative Flow

N/A

2.4.5.8 Post-conditions

CODECO analyses periodically the status of the overall system (data-compute-network infrastructure) continuously proposing adjustments.

Data plane continues without interruption.

2.4.5.9 High Level Illustration

See **Figure 35**.

2.4.5.10 Potential Requirements

2.4.5.10.1 Deployment KPIs:

Time to completion of a task (latency): reduction in 20% due to decentralized control.

10% of reduction in the number of collisions.

10% improvement of energy efficiency of the network (overall involved nodes) and eventual network lifetime

Number of nodes supported (at least 5); reduction in failures (resilience).

2.4.5.10.2 Non-functional requirements:

Scalability, capability of CODECO to cope with an increasing number of application deployments across variable fleet sizes (large, heterogeneous) and towards mobile devices.

Privacy, capability of CODECO to manage a varying infrastructure across a single or different locations (multi-domain environments).

Resilience and availability, capability of CODECO to support five nines system availability in the verge of network interference and intermittent connectivity.

2.4.6 CODECO P6: Automated Crownstone Application Deployment for Smart Buildings

Contact: ALMENDE; Andries Stam (andries@almende.org)

2.4.6.1 Description

In CODECO P6, we will focus on novel mechanisms for automated deployments of smart office/smart building applications on the Crownstone Platform. In this context, an application is defined as a collection of related functionalities realized by means of a set of interconnected application components which can run either in the Cloud, on the Crownstone Hub, or inside a Crownstone Node. The key issue we will address is how the CODECO technologies can help with automated deployment of multiple applications on the Crownstone platform, both in single cluster situations (where multiple Crownstone Hubs form a single manageable entity with a single user base), and in multi-cluster situations (where multiple Crownstone Hubs form multiple manageable entities with different (but potentially overlapping!) user bases).

The Crownstone platform technology has been developed within the Almende group during the past years. The five main constituents of the technology are:

1. Smart lustre terminals called **Crownstone nodes**, which can be mounted inside power outlets. Each Crownstone node has five capabilities: switching on and off (or dimming) the devices attached to the power outlet, measuring the power consumption of the device attached to the power outlet, maintaining BLE connections with wireless sensors and/or actuators, communicating with other Crownstone nodes via Bluetooth mesh, and running small apps called Microapps on the processor inside the Crownstone node.

2. USB-sticks called **Crownstone Bridges**, which are Crownstone nodes with their UART connected to USB male socket, but without the technology to switch on/off devices and measuring power consumption. Bridges are used to connect a Bluetooth Mesh network to a Crownstone Hub.

3. Raspberry Pis with a specific software stack installed called **Crownstone Hubs**. These are used to collect data from larger collections (at most 256) of Crownstone nodes, to process data, to deploy Microapps to Crownstone nodes, and to connect to the Cloud.

4. A Cloud Service called the **Crownstone Cloud,** which serves to administer and exchange information about spheres (i.e., buildings / environments with Crownstone nodes that are connected to each other), rooms, Crownstone nodes, smartphones, and their relationships.

5. A React Native based app called the **Crownstone App**, which lets your smartphone act as a beacon which can be detected by Crownstone nodes and provide a management dashboard for managing all details of the Crownstone platform within a single sphere.

The Crownstone platform was originally developed as a universal domotics solution for the consumer market. However, recently, Almende has decided to make a shift from B2C to the B2B market, offering the Crownstone platform as a universal smart building technology for office environments, industrial environments, etc. This poses completely new challenges on the technology, which are partially addressed in this use case.

2.4.6.2 Source

2.4.6.3 Roles and Actors

Actors:

Crownstone application developers

Building managers

Roles:

A microapp developer develops a microapp to be run on subset of Crownstone in the network, and wants to deploy apps quickly without moving within Bluetooth range to Crownstone, copy-pasting MAC addresses, etc. They set relevant parameters in config file and the microapp deployment manager will make sure the microapps are correctly uploaded to the designated Crownstone.

A building manager wants to deploy applications without worrying about the network topology of Crownstone. They can push a single configuration file and FADO manager will orchestrate the deployment of the applications among the hubs. Any issues around deployment (e.g., resource problems) are fed back to building manager, if possible, with directions for fixing the issues.

A summary of the business impact is as follows:

The management and deployment of crownstones is currently highly centred around the use by consumers. This entails the use of the Crownstone app to manage crownstones, locations and schedules including constraints.

A business-oriented management scenario should enable managing organisations to organise sets of crownstones into spaces, users and (micro)applications. A typical scenario might require the sharing of all users over all buildings, while a set of crownstones cannot currently exceed 256 elements. The option to have multiple sets of crownstones recognising and tracking most if not all users(/assets) is a much-requested feature that will increase the target market.

Our crownstone platform enables the deployment of multiple different applications, developed by multiple third parties, on a single infrastructure inside a smart building, thereby avoiding that every new end user application installed in the building always introduces new hardware components to run the application on. The difficulty, however, is that we need to have better tools to manage and monitor the portfolio of installed applications in a single building or even over multiple buildings, as they share the same infrastructure. Our new tooling will build these tools to overcome this new management problem.

2.14.6.4 Pre-conditions

Medium-to-large-scale buildings with a multitude of rooms equipped with sensors and/or actuators.
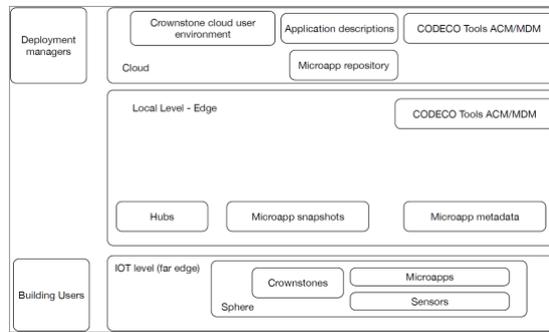
Bluetooth mesh topology between crownstones and hubs

2.4.6.5 Triggers

New Crownstone setup across a building or room required.

2.4.6.6 Normal Flow

The Microapp deployment architecture illustrated in **Figure 36** can be divided in a Cloud level, a local (per building/apartment) level and AIoT level.

Figure 36: CODECO P6 system architecture

Crownstone are already administered in spheres. A sphere is a collection of crownstones with associated locations (e.g., rooms) and users. A digital twin of this representation is present in the crownstone Cloud.

Microapps can communicate with the crownstones and the local hub, that serves as a relay for the Cloud-based environment. Collections of microapps can be configured (together with optional added sensors) in the Cloud and adapt to the locally available resources.

The Cloud environment will be combined with the CODECO ACM/MDM functionalities to allow for this adaptive distribution and deployment of microapps. The crownstone firmware is already able to receive the microapps. The CODECO MDM functionality will be made available on the hub level to execute the configurations defined at the Cloud level.

The initial UML use-case diagram is provided in **Figure 37**.

Figure 37: P6 UML Use-case Diagram

2.4.6.7 Alternative Flow

N/A

2.4.6.8 Post-conditions

Once the new infrastructure is active, measurements can be performed in the Crownstone mesh.

2.4.6.9 High Level Illustration

See **Figure 36** and **Figure 37**.

2.4.6.10 Potential Requirements

2.4.6.10.1 Deployment KPIs:

Number of deployed Crownstone nodes over time, with CODECO, in comparison to regular operation.

Network reliability over time (K8s as baseline for deployment).

Latency

Throughput

Energy consumption

2.4.6.10.2 Non-functional requirements:

Non-invasiveness (e.g., Crownstone basic functionality always remains).

Privacy-preserving with respect to data captured by sensors related to people in the building and their behaviour.

2.4.7 COGNIT - CCAM services enhanced with AI assisted CEI framework

2.4.7.1 Description

Our use case focuses on Connected, Cooperative, and Automated Mobility (CCAM) services within the Cloud-Edge Continuum, aiming to enable safer, more efficient, and intelligent mobility through distributed computing and seamless data sharing among vehicles, edge devices, and the cloud.

Specifically, we have implemented the Public Transit Service (PTS) and the Time-to-Green (TTG) service across 114 regulated intersections in the city of Granada (Spain). A total of 38 Mobility-Hub (ACISA's Traffic light Controller) with C-V2X/DSRC capabilities have been installed to manage those intersections.

One public transit line (Line 4) has been equipped with homologated C-V2X/DSRC On-Board Units (OBUs) on all buses. Thanks to the PTS service, these buses can request priority passage at equipped intersections, effectively reducing overall transit time.

Additionally, other vehicles equipped with C-V2X OBUs may receive real-time information about the remaining time until the green light phase, enhancing driving experience.

Public bus priority systems are commonly used in the transport management of cities, but one of the key advantages of using standardized V2X equipment is that it can be reused to provide more than 30 different use cases harmonized for interoperability by the C-Roads Platform[5].

Those intersections in Granada will be gradually updated with additional V2X services using the same infrastructure.

---

[5] https://www.c-roads.eu/platform.html

Why this Use Case?

There's a push toward safer and smarter road transport via vehicle automation and connectivity. CCAM is being promoted across Europe (e.g., through EU-funded initiatives) as a solution to reduce accidents, improve traffic flow, and cut emissions.

Traditional centralized systems can't handle the low latency and real-time processing needed for CCAM. This use case leverages the COGNIT cloud-edge continuum framework to:

- Distribute computation close to the data source (e.g., in-vehicle or roadside edge nodes).

- Seamless resource provisioning by far edge devices through Function as a Service (FaaS) paradigm in the continuum.

- Offload intensive tasks to Edge servers or cloud for more complex processing and learning (Intersection Digital Twin simulation functionality).

- An AI based Orchestrator provided by the COGNIT framework decides where to provision resources to execute the requested FaaS.

It is foreseeable that many CCAM services will be allocated along 5G network slices[6]. 5G network slicing allows for the creation of multiple virtual networks on a shared physical infrastructure, each tailored to meet specific performance, capacity, and functionality requirements. This capability is particularly beneficial for CCAM services, which demand diverse and stringent network characteristics such as ultra-reliable low latency and, eventually, high bandwidth. As 5G technology matures, network slicing is transitioning from a theoretical concept to practical implementation across various sectors[7] such as manufacturing, healthcare, and emergency services, demonstrating its versatility and effectiveness.

Given these advancements, it is anticipated that CCAM services will increasingly leverage 5G network slicing to ensure the required Quality of Service (QoS). This integration will facilitate the deployment of intelligent transportation systems, enhance vehicular communication, and support the development of autonomous driving technologies.

2.4.7.2 Source

COGNIT (A Cognitive Serverless Framework for the Cloud-Edge Continuum), H2020 European project (https://cognit.sovereignedge.eu/)

2.4.7.3 Roles and Actors

The COGNIT project involves a diverse set of roles and actors collaborating to develop an AI-enabled cognitive cloud for Europe's cloud-edge continuum. Below is an overview of these roles, their responsibilities, relationships, and associated actors:

- **Roles in the Use Cases:**

  - **End Users:** Individuals or organizations utilizing the applications developed within the COGNIT framework, such as city residents, energy consumers, or entities responsible for wildfire detection.

  - **Vertical Industry Partners:** Organizations operating within specific sectors that implement and benefit from the COGNIT framework. For instance:

    o Smart Cities: ACISA coordinates this use case, focusing on urban CCAM services enhanced with continuum resources.

    o Wildfire Detection: Nature 4.0 leads efforts in environmental monitoring and disaster prevention.

---

[6] https://blog.sasken.com/5g-network-slicing
[7] https://elpais.com/proyecto-tendencias/2025-01-31/network-slicing-de-concepto-teorico-a-una-realidad-gracias-a-la-red-5g.html

- o <u>Energy</u>: Phoenix Systems and Atende Industries collaborate on energy management solutions.
- o <u>Cybersecurity</u>: CETIC and SUSE address security aspects within the cloud-edge continuum.

- **Communication Network Providers/Operators:** Entities supplying the necessary networking infrastructure to support data transmission between edge devices and the cloud.

- **IoT Device Manufacturers:** Companies producing sensors, actuators, and other hardware components deployed in various environments to collect data.

- **IoT Platform Providers:** Organizations offering platforms that facilitate the integration, management, and analysis of data from IoT devices within the COGNIT framework, like ACISA or Phoenix.

- **Cloud Edge Service Providers:** Entities providing cloud computing resources where certain computational tasks are executed as part of the serverless framework, such as OpenNebula.

- **Research Institutions:** Academic and research organizations contributing to the development and validation of the COGNIT framework through scientific research and innovation, like Ümea University, CETIC or RISE.

Relationships Between Roles:

- **End Users** interact with applications developed by vertical Industry Partners, utilizing data processed through the COGNIT framework.

- **Vertical Industry Partners** collaborate with IoT Device Manufacturers to deploy necessary hardware and with IoT Platform Providers to manage and analyze data.

- **Communication Network Providers/Operators** ensure reliable connectivity between IoT devices, edge nodes, and cloud services, facilitating seamless data flow.

- **Cloud Service Providers** offer computational resources for tasks that cannot be handled at the edge, working in tandem with edge computing resources to optimize performance.

- **Research Institutions** support all stakeholders by providing insights, developing innovative solutions, and validating the framework's effectiveness across different use cases.

Actors and Their Roles:

- **Open Nebula**: Act as a project coordinator and technical lead. The initial deployments of the COGNIT framework leverages Open Nebula IaaS solutions.

- **Ikerlan**: Act as a lead development of a distributed Function-as-a-Service (FaaS) paradigm. This paradigm aims to enable Internet of Things (IoT) and edge devices to offer compute-intensive applications through intelligent task offloading to the cloud-edge continuum

- **Ümea University, CETIC, RISE:** Collaborate as research institutions across various roles to provide scientific expertise, such as AI and cybersecurity, and contribute to framework development, and validate use cases.

- **ACISA:** Acts as the Vertical Industry Partner for the Smart Cities use case, coordinating efforts to integrate the COGNIT framework into urban mobility environments.

- **Nature 4.0:** Serves as the Vertical Industry Partner for the Wildfire Detection use case, focusing on environmental monitoring applications.

- **Phoenix Systems and Atende Industries:** Function as Vertical Industry Partners in the Energy use case, developing smart energy management solutions.

- **CETIC and SUSE:** Operate as Vertical Industry Partners for the Cybersecurity use case, addressing security challenges within the cloud-edge continuum.

2.4.7.4 Pre-conditions

COGNIT is a RIA project that started at **TRL 2** and aims to achieve **TRL 5** by the project's conclusion. It will reach a higher TRL with the lead of Open Nebula and through further collaboration of other COGNIT partners and the Open-Source community.

2.4.7.5 Triggers

The trigger is the function sent to the Cognit framework in the form of FaaS. In the case of Smart City use case, the trigger is the request of transit priority by a public bus.

2.4.7.6 Normal Flow

In the Smart city use case developed by ACISA, this is a normal flow of exchanged data between key entities:

1. **Vehicles:** Equipped with On-Board Units (OBUs), vehicles send and receive data through V2X communication. Public buses may send priority request to M-Hub to reduce transit time at intersections. They may also receive information like traffic light timings, intersection layout, or information.

2. **Infrastructure:** Roadside Units (RSUs) and Traffic Light Controllers (M-Hubs) are part of the infrastructure. They communicate with vehicles, providing information about traffic conditions, signal timings, and receiving requests (e.g., for traffic signal priority). M-Hubs also exchange data with the COGNIT framework through FaaS paradigm, and Saturno, ACISA's Smart Mobility Suite.

3. **IoT Platform:** The IoT platform (like ACISA's Saturno) collects and manages data from various sources, including M-Hubs, other traffic subsystems or external information systems. It may also provide data and services to other entities, like providing historical and real-time traffic data to Digital Twins.

4. **COGNIT Framework:** This framework manages the cloud-edge continuum, handling the deployment and orchestration of applications and services. It receives function calls and data from entities like Saturno, and provides resources and services in return, such as the orchestration of resources to execute the computation demanded in FaaS. In our use case, it manages the execution of traffic simulations to assit intersection Digital Twins.

5. **Digital Twins:** Acisa implements a distributed Digital Twins of each intersection to better evaluate its status and evolution. It receives data from various sources, including vehicular data, traffic information, and environmental data, to create a digital representation of the physical environment. They may also send data or commands to infrastructure, such as adjusting traffic signal timings.

Each time a priority request reaches the M-Hub, an evaluation is performed with the aim of its Digital Twin to determine whether it is appropriate to grant priority or not. Priority will be granted if the process does not significantly disrupt the intersection flow and if it improves the bus's transit time according to simulations results (see **Figure 38**)

2.4.7.7 Alternative Flow

2.4.7.8 Post-conditions

When a priority request reaches the M-Hub, its Digital Twin evaluates whether to grant it. Priority is given if it minimally disrupts intersection flow and improves bus transit time based on simulation results.

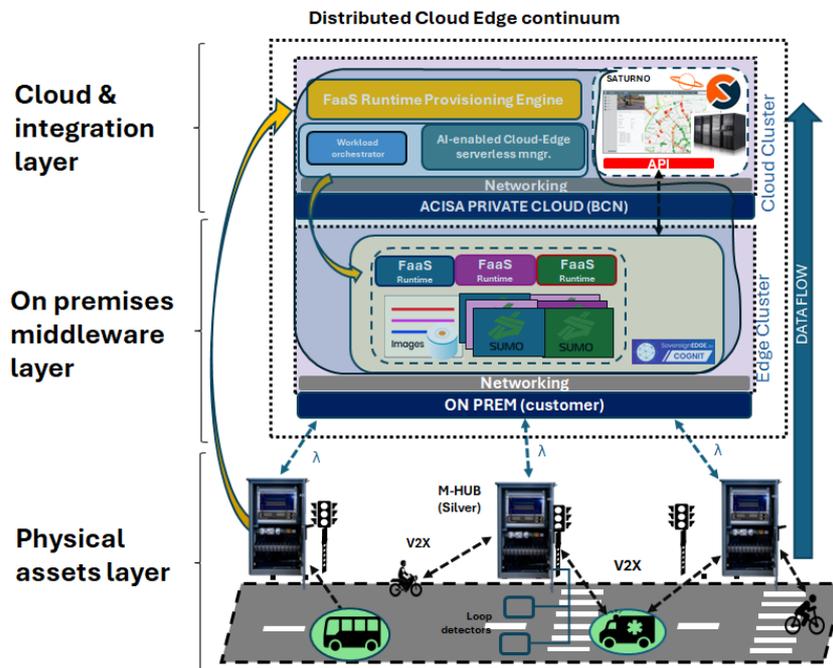### 2.4.7.9 High Level Illustration

Figure 38: COGNIT Smart City use case. Priority request flow representation.

2.4.7.10 Potential Requirements

Current Smart City use case is deployed using C-V2X communication standard (3GPP PC5 Mode 4), where OBU and RSU also communicate through 4G for monitoring and maintenance services. PC5 interface was released in 3GPP release 14 and it provides direct communication between the vehicle OBU and the infrastructure RSU without the need of any cellular network.

The potential requirements consider the future evolution C-V2X communications within the 3GPP standardization body, NR-V2X, in releases 16+.

Functional Requirements

- Roundtrip latency under 2ms.

- Secured communication through cryptography methods.

- The network must support millions of connected vehicles in cities, that must maintain synchronization, and transparent handover as vehicles move along the city.

- Eventually, V2X services might be allocated in dedicated slices within the 5G network infrastructure.

- CCAM-related services might be classified in different 5G service categories:

  o Safety use cases (real-time) → URLLC

  o Entertainment/internet in the car → eMBB

  o Vehicle sensors, smart traffic systems monitoring and diagnostics → mMTC

Non-Functional Requirements.

- Interoperability between car manufacturers, network operators and infrastructure.

- Standard-based communication based on ETSI V2X standards.

- Guaranteed and programmatically configured Quality of Service (QoS)

- Isolation from other network traffic

- Scalable communication between systems interconnects massive volumes of vehicles with the road infrastructure and other vehicles through direct link (PC5 interface) or through the 5G network (Uu interface).

- Reliable communication between vehicles and infrastructure. There are CCAM services related to safety and security (Day 1, Day 1.5).

2.4.7.11 Radio Specific requirements

2.4.7.11.1 Radio cell range

- Specification of expected maximum and typical radio ranges (indicate if LOS/NoLOS):
  - NoLOS
- Does the radio link crosses public spaces? Or is it constrained to indoor or customer premises?
  - Radio link must operate in crowded city outdoors
- Is Multicell required?
- (If YES, specify the required scope of the multicell arrangement. I.e. "building", "city", "global"):
  - Probably yes, City scope.
- Is handover required? Seamless? Tolerable impact in delay and jitter?
  - Yes, seamless.
- Mobility: maximum relative speed of UE/FP peers
  - We may consider 80-90 Km/h in urban scenarios.
- Special coverage needs: I.e maritime, aerial
  - Urban and interurban areas coverage

2.4.7.11.2 Bandwidth requirements

2.4.7.11.3 URLLC requirements

- **Required Latency**
  - 2 ms roundtrip
- **Required Reliability**
  - Maximum tolerable jitter

2.4.7.11.4 Radio regimens requirements

- Desired and acceptable radio regimens
- (describe the desired and acceptable radio regimens: I.e.: licensed - public mobile, licensed – specific license, license-exempt)
  - licensed - public mobile

2.4.7.11.5 Other requirements

- UE power consumption
- Rechargeable or primary battery?
  - V2X systems are normally powered by DC current.
- Acceptable battery life
  - N/A
- Is terminal location required? location accuracy?
  - Location accuracy is relevant in V2X services. < 1m precision is desirable.

2.5 Smart Agriculture

2.5.1 COMMECT: Monitoring of Pest Insect Traps

2.5.1.1 Description

Olive tree is a plant native to the Anatolian region, and its main products (olive fruit and olive oil) have been considered important food and commercial products since ancient times. In recent years, with the increasing interest in healthy life and nutrition in the world, the importance of the production and consumption of table olives and olive oil is increasing. Türkiye constitutes a vital gene pool with 93 domestic olive varieties. Modern planting systems, mechanization and digitalization are taking place rapidly in olive farming around the world. On the other hand, Türkiye is not at the same level as developed countries in terms of integrating technology and digitalization into olive farming. The biggest reason for this is the low-income level of olive farming producers and the lack of telecommunication coverage in the countryside. Lack of agriculture knowledge of the olive producer, difficulty accessing digital products, and technological infrastructure can be counted as other reasons. Experts' training of rural people, eliminating the deficiencies and mistakes in traditional olive farming, is expected to positively affect the yield and quality and contribute to Türkiye's economy. Our main intention is to make the right timely decisions to produce higher-yielding quality olives via monitoring the environmental and pest population to better plan their olive orchard activities.

2.5.1.2 Source

[Bridging the digital divide and addressing the need of Rural Communities with Cost-effective and Environmental-Friendly Connectivity Solution (COMMECT) HE project](#)

2.5.1.3 Roles and Actors

Olive farmers who undertake the daily work and management of the olive orchard, are mentioned as one group of essential stakeholders. This group is more vulnerable and uninformed about the effects of climate change, disease outbreaks, and economic changes and uncertainties.

Olive processing industrialists in the value chain after olives have been harvested. This group also comprises representatives from cooperatives and olive production unions.

Olive machine producers who produce process machine for table olive and oil olive for the farmers and industrialists.

2.5.1.4 Pre-condition

The main pre-condition is to live with the potential risk of monitoring pest insect traps that could increase the insect population surrounding orchards that do not use these traps.

2.5.1.5 Triggers

The agricultural sector has become increasingly reliant on technological advancements in recent years. The Internet of Things is a new technology expected to boost productivity in farming and agricultural activities, leading to higher yields and lower costs per kilogram of olives. Olive agriculture worldwide rapidly adopts modern planting systems, mechanization, and digitalization.

2.5.1.6 Normal Flow

Commonly, the steps are the follows:

1. Olive fruit fly is the primary pest of olive orchards and causes a significant amount of yield and quality losses. The larvae of olive fruit flies give rise to direct damage.

2. Larvae from the egg eat the fruit flesh by opening galleries around the seed. Its damage in the olive sector is essential since it causes the fruits to decay and the acidity in the olive oil to rise.

3. Pesticides are applied when necessary and on time. This is possible by monitoring the olive fly adults in nature.

4. Spraying should be done before the olive fly becomes harmful by using digital traps to identify the first flies, such that when the number and species of pests exceed a certain threshold.
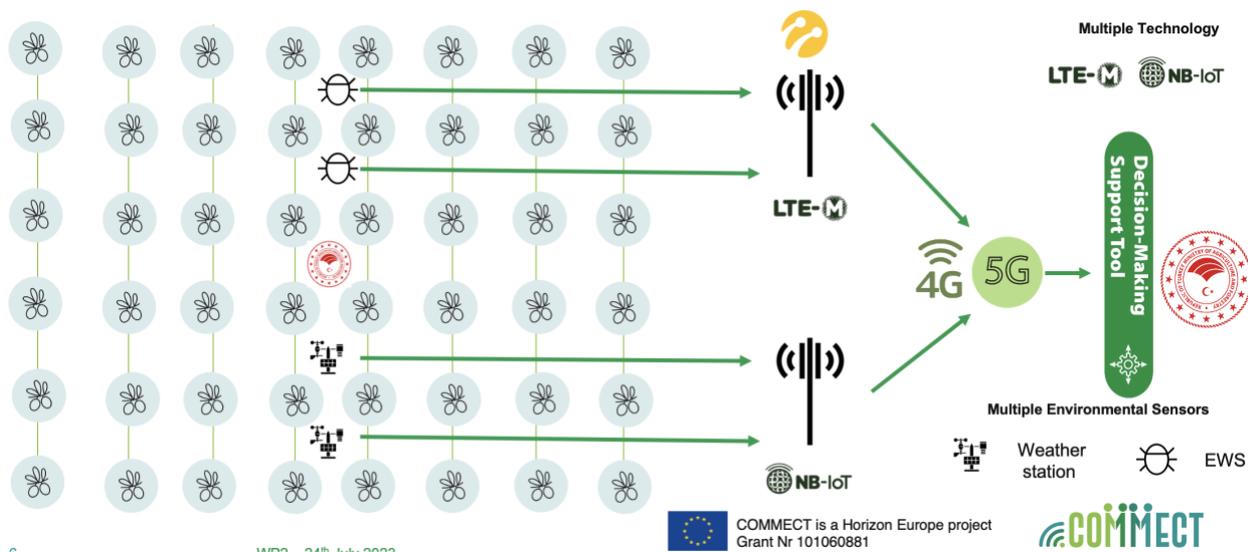
2.5.1.7 Alternative Flow

N/A

2.5.1.8 Post-conditions

Once the risks have been minimized or solved, the olive pest insect traps inform the farmers and industrialists about the population of the insects.

2.5.1.9 High Level Illustration

2.5.1.10   Potential Requirements

The following requirements have been defined for this use case:

Fly Detection Accuracy: potential risk analysis based on olive fruit fly population

Uplink Throughput: uploading of machine vision pest monitoring photo

Power Consumption Decrease: extended battery durability of sensor equipment

2.5.1.11   Radio Specific requirements

Does the radio link crosses public spaces? Or is it constrained to indoor or customer premises?: The base stations give services to the traps and, at the same time, the mobile network operator customers in the region, like rural or suburban areas. The side to side distance is approximately between 5-10 km in rural or suburban areas

Mobility: No mobility is needed

2.5.1.11.2   Bandwidth requirements

Peak data rate: 4 Mbit/s

Average data rate: 2 Mbit/s

Is traffic packet mode or circuit mode? Packet mode

2.5.1.11.3   URLLC requirements

N/A

2.5.1.13.4 Radio regimens requirements

Desired and acceptable radio regimens: Licensed – public mobile.

2.5.1.11.5   Other requirements

UE power consumption

Rechargeable or primary battery? Rechargeable battery

Acceptable battery life: 10 years

Is terminal location required? location accuracy? No, it is not needed.

2.5.2   COMMECT: Securing crops and equipment

2.5.2.1 Description

Use Case titled "Securing Crops and Equipment," is implemented in the context of the COMMECT project and its Living Lab Serbia initiative. It aligns with the increasing demand for digitalization in agriculture focusing on deploying edge ML computing to address challenges in agriculture, such as securing assets from theft, monitoring growth of the crops, and supporting improvement of the overall operational efficiency of agriculture activities in the field.

Implementation of the use case involves several key steps. First, video cameras and edge ML devices are installed in the field, powered by renewable energy sources with battery backup to ensure sustainable operation. These devices are configured and managed remotely, allowing for energy-efficient adjustments and the customization of machine learning models to meet field-specific requirements. Data is then continuously captured and video streams analysed to get insights into the growth of the crops as well as to identify suspicious activities by detecting people and vehicles in the field. , and. The performance of the edge devices is monitored to balance processing power and energy use. Remote management capabilities further enhance the adaptability and reliability of the solution enabling dynamic changes of AI algorithms based on the needs of agricultural operations.

Living Lab Serbia focuses on advancing sustainable agriculture and preserving the natural environment in Gospođinci village, located in the Vojvodina province of northern Serbia. This initiative encompasses five diverse use cases, each addressing critical challenges in rural development. Through these use cases, Living Lab Serbia provides practical demonstrations of how cutting-edge digital technologies can empower local communities, optimize agricultural practices, and protect biodiversity.

2.5.2.2 Source

[Bridging the digital divide and addressing the need of Rural Communities with Cost-effective and Environmental-Friendly Connectivity Solution (COMMECT) HE project](#)

2.5.2.3 Roles and Actors

Use case "Securing crops and equipment" involves several key actors, each playing a vital role in its implementation and success:

Farmers and Farmer Associations are the primary end users and beneficiaries. They adopt improved agricultural practices, integrate digital solutions, and collaborate in data sharing to optimize farming operations.

Companies (e.g., digital service providers, hardware manufacturers, solar trailer producers) are responsible for providing power and network infrastructure along with IoT devices, deploying solutions, and training farmers to use technology effectively.

These companies also utilize pilot sites to test and refine their products and services, promote sustainable practices, and foster knowledge exchange and community collaboration.

Local government entities act as policy enablers and regional development supporters. They establish agricultural policies and regulations, attract investments, and encourage sustainable development while working to preserve natural habitats and ecosystems.
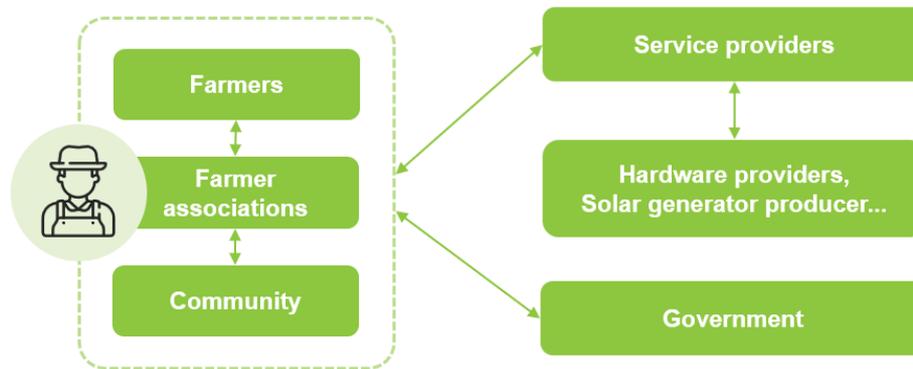
2.5.2.4 Pre-conditions

The successful implementation depends on meeting the following conditions:

Technical pre-conditions

Infrastructure readiness: Edge ML devices with sufficient processing power and reliable power supplies must be available to support operations in remote areas.

Connectivity: A stable communication network (e.g., LoRa, 4G/5G, or Wi-Fi) is essential for transmitting alerts and data securely using robust protocols.

Device setup: Cameras must be installed and calibrated to capture accurate video data, covering key monitoring areas.

ML model deployment: Pre-trained ML models optimized for detecting people, vehicles, and audio patterns should be deployed and configured on edge devices.

Non-technical pre-conditions

End-user training: Farmers and users must be trained to operate the system, interpret data, and respond to alerts effectively.

Stakeholder engagement: Collaboration among farmers and technology providers.

2.5.2.5 Triggers

Specific triggers are necessary to initiate actions such as analyses and alerts, based on real-time data processing:

Detection of movement: When people or vehicles are detected in the monitored area, the system analyzes video feeds to identify activity and sends picture with captured object and alert (e.g., SMS or app notification).

Crop monitoring events: Significant changes in crop growth parameters (height, size, or signs of stress) and activity tracking (spraying, plowing, sowing) providing actionable insights for job management.

2.5.2.6 Normal Flow

The typical operation involves seamless data exchange between key components, following these steps:

Continuous data collection: Cameras deployed on edge devices, powered by a mobile solar generator with battery backup, capture real-time video streams from the monitored area.

Real-time data processing: Edge computing devices analyze collected data using ML algorithms, detecting and counting people or vehicles and monitoring crop growth. Each device processes at least two video streams simultaneously.

Event detection: Potential threats or events, such as unauthorized vehicles or theft attempts, are identified by classifying anomalies based on trained ML models.

Communication with IoT Platform: Data is transmitted from the edge devices through switch to a router which forwards data to the IoT platform and agroNET solution via a 4G network.

User notifications: Alerts and insights are sent to farmers via mobile or web applications, providing information on detected events.

Remote monitoring and configuration: configuration of edge devices, updates and changes of ML algorithms, monitoring power consumption, and optimizing battery usage.

2.5.2.7 Alternative Flow

N/A

2.5.2.8 Post-conditions

Once the use case is deployed and the system is fully operational, farmers will have a solution for monitoring and securing crops and equipment. The system will notify farmers when anomalies are detected, sending pictures and relevant data to the agriculture platform for further analysis and action.

2.5.2.9 High Level Illustration

The system architecture integrates this use case with others implemented within the Living Lab, as illustrated in the figure below.
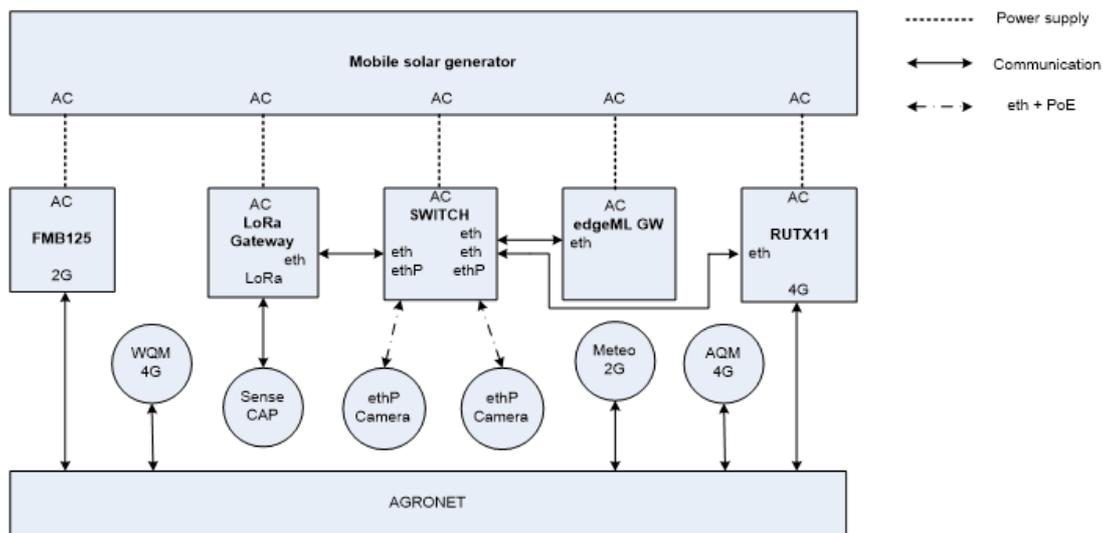
2.5.2.10   Potential Requirements

Technical requirements:

Video processing: The edge computing infrastructure must support real-time processing of video streams. The insights generated can be variable, depending on the needs of particular agricultural operations and deployment context. A minimum of two video streams should be processed simultaneously.

Edge ML remote configuration: power consumption, battery status, and power requirements of active ML algorithms must be monitored. Edge devices should be remotely configurable to optimize power consumption as needed.

The table below summarizes technical requirements for this use case. Table is copied from *Deliverable 1.2 Report on COMMECT requirements and KPIs*.

Requirement 3. ID	4. Description	5. Technical Requirement(s)	6. Target Value

| R5.8 | Edge ML computing infrastructure | Video processing | >=2 |
| R5.9 | Power consumption and power requirements monitoring | Edge ML remote configuration | Remote configuration supported and enabled. |

Non-functional requirements:

Flexibility: The system must be adaptable to various agricultural environments, crop types, and evolving technology requirements.

Integration capability: The system should seamlessly integrate with existing technologies and digital IoT solutions.

Continuous operation: The system must ensure continuous operation to maintain effective security and crop monitoring functions.

2.5.2.11 Radio Specific requirements

2.5.2.11.1 Radio Coverage

Outdoor radio link

Low power, long-range for sensors in the field

WiFi or cable for video cameras

Mobility is welcome, but not mandatory

2.5.2.11.2 Bandwidth requirements

The solution addressing the use case is designed to maximize the edge processing and minimize the amount of data transferred to cloud. Current 4G connectivity supports well the scenario. However, if edge processing is not used and more remote management required (e.g., remote control of drones monitoring and spraying crops), that increased throughput and minimal latency would be required.

2.5.2.11.3 URLLC requirements

N/A

2.5.2.11.4 Radio regimens requirements

No particular requirements. What is important, is availability of reliable communication network.

2.5.2.11.5 Other requirements

Sensors: long duration battery (at least one season), replaceable or rechargeable.

Edge: rechargeable without interruption.

Ability to automatically determine location of the edge server is preferrable. GPS accuracy. Location of sensors is welcome if it does not significantly impact the battery lifetime.

3.   Emerging Topics

This section describes emerging topics that are related to IoT and Edge Computing and can impact the specifications and deployments of 5G. Those emerging topics are:

1. Digital Twin (DT)

2. Deep Edge, Terminal and IoT Device Integration in B5G communication infrastructure

3. Edge, Mobile Edge Computing and Processing

4. Network and Server security for edge and IoT

5. Plug and Play Integrated Satellite and Terrestrial Networks

6. Autonomous and Hyper-connected On-demand Urban Transportation

7. Opportunities for IoT Components and Devices

8. EU legislative framework.

3.1 Digital Twin

It is important to define the meaning of Digital Twin (DTw) concept before proceeding, as it has been interpreted in many ways in the past years. It is important to have a common understanding what are implication of such concept and, more, to properly address possible impact and benefits of this approach considering adoption of 5G.

The Digital Twin in its original form is described as a digital informational construct about a physical system, created as an entity on its own and linked with the physical system in question. One of the first domain it was adopted was in Aerospace Industry, where it was referred as "To address the shortcomings of conventional approaches, a fundamental paradigm shift is needed. This paradigm shift, the Digital Twin, integrates ultra-high fidelity simulation with the vehicles on-board integrated vehicle health management system, maintenance history and all available historical and fleet data to mirror the life of its flying twin and enable unprecedented levels of safety and reliability." [TaQi19].

In such perspective the key aspect referred to DTw is the accurate representation of the structure, the status and the actual behaviour of a physical object in term of collection of relative data. The most relevant aspect is in such way associated to be able to collect in "proper" way enough and with adequate granularity information or in other words Digital Twin in its origin describes a product mirroring its available informational status.

Based on the given definitions of a Digital Twin an evolution took place to represent increased capacity of DTw to provide enriching services based on embedded technologies able to structure, elaborate and forecast the information related to the physical object. So, in manufacturing domain, one new definition can be adopted to better describe this aspects. "The DT consists of a virtual representation of a production system that is able to run on different simulation disciplines that is characterized by the synchronization between the virtual and real system, thanks to sensed data and connected smart devices, mathematical models and real time data elaboration. The topical role within Industry 4.0 manufacturing systems is to exploit these features to forecast and optimize the behaviour of the production system at each life cycle phase in real time." [TaCa19].

A relevant aspect that needs to be considered is now the way the DTw interact with the physical world, in fact we have for sure the need to gather information to "build" the basic content of the digital twin, but other important questions emerge:

1. Data collection is carried out manually or automatically?

2. Data collection is executed only once at the creation of the DTw or carries on for its entire life?

3. Internal representation of the physical object is static or is dynamically updated?

4. Any possible result of DTw elaboration can be "returned" to the Physical object to improve its behaviour (efficiency, safety, duration,) or to a third entity to provide any value?

Before answering in full to these questions, let first focus on the interactions between Physical Object and DTw. We introduce this terminology for DTw, as digital counterparts of physical objects. We consider these definitions: Digital Model, Digital Shadow and Digital Twin strictly speaking, see [Glaes12].

A Digital Model is a digital representation of an existing or planned physical object that does not use any form of automated data exchange between the physical object and the digital object.
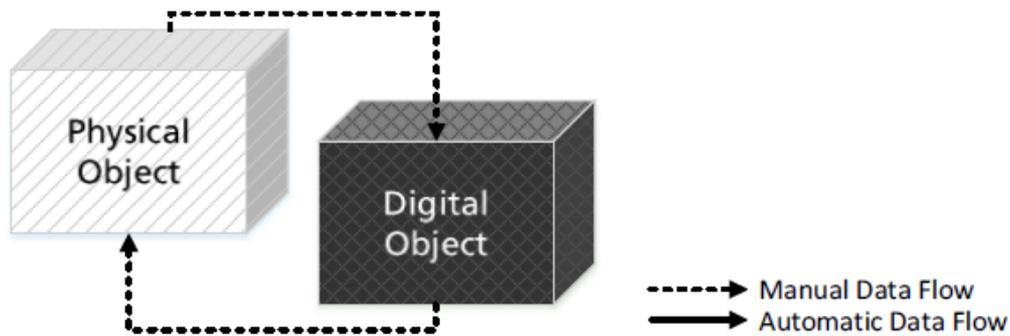

Figure 39: Data Flow in a Digital Model

Based on the definition of a Digital Model, if there further exists an automated one-way data flow between the state of an existing physical object and a digital object, one might refer to such a combination as Digital Shadow.
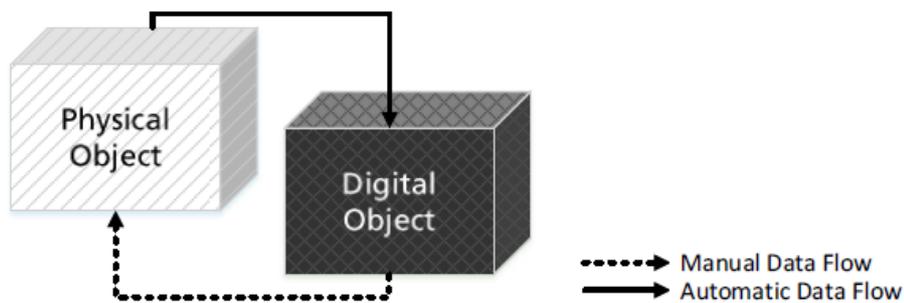
Figure 40: Data Flow in a Digital Shadow

If further, the data flows between an existing physical object and a digital object are fully integrated in both directions, one might refer to it as Digital Twin.
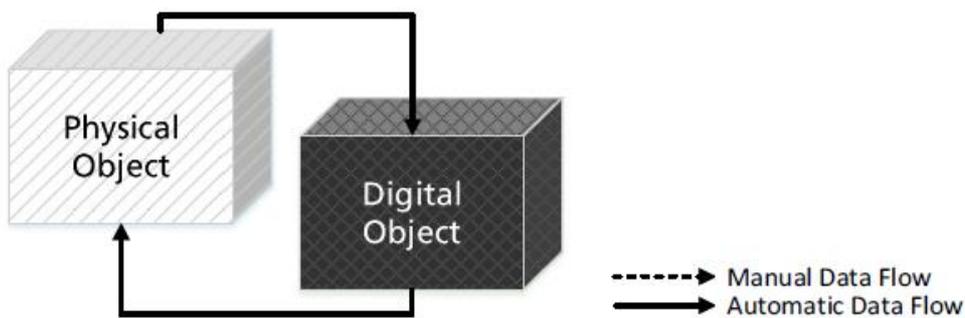
Figure 41: Flow in a Digital Twin

131

A more structured representation of DTw that encompasses an advanced bi-directional information flow between physical and digital entity and internal capacity able to elaborate and enrich information including capability to provide added value or services.

We can represent it with the following representation in **Figure 42**, see [GaRo12].
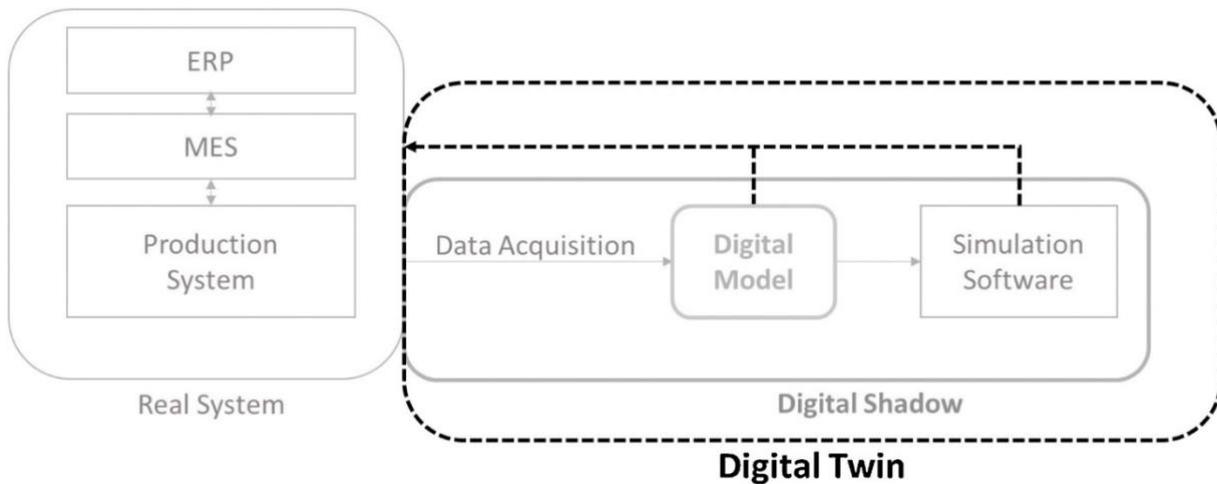
Figure 42: Digital Twin (DT) schema, copied from [GaRo12]

State-of-the-art technologies such as the IoT, Wireless and Mobile Communication, cloud computing, big data analytics, and AI have greatly stimulated the development of smart manufacturing. An important prerequisite for smart manufacturing is cyber–physical integration, which is increasingly being embraced by manufacturers. As the preferred means of such integration, CPS and digital twins have gained extensive attention from researchers and practitioners in industry, see [KrKa18]. The essence of CPS is to add new capabilities to physical systems using computation and communication, which intensively interact with the physical processes and, if needed, is able to involve as part of the process also human operators and/or decision makers, providing added value services all along the lifecycle of the production process and eventually of the product.
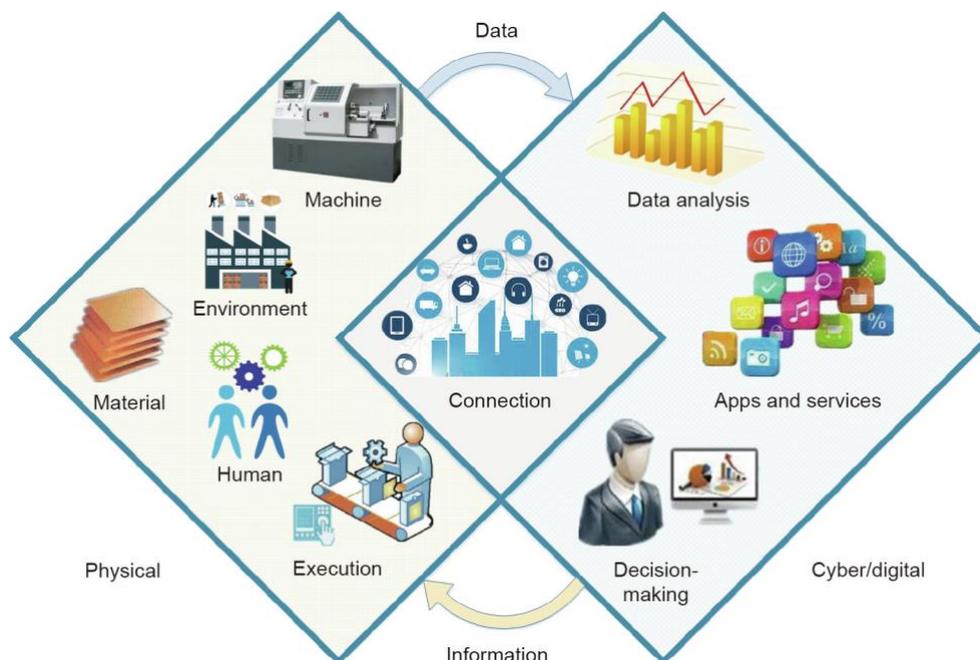
Figure 43: Mapping between physical and cyber/digital worlds, copied from [KrKa18]

CPS Cyber Physical concept as evolution of the Digital Twin is at the base of new paradigm, as Industry 4.0 in Manufacturing, Logistics and Operation. In **Table 24** the differences between the two terms are formalized.

Table 24: Correlation and comparison of CPS and DTs, copied from [KrKa18]

Table 1
Correlation and comparison of CPS and DTs.

Items	CPS	DTs
Origin	Coined by Helen Gill at the NSF around 2006	Presented by Michael Grieves in a presentation on PLM in 2003
Development	Industry 4.0 listed CPS as its core	Not much attention paid to DTs until 2012
Category	Akin to a scientific category	Akin to an engineering category
Composition	The physical world and the cyber world, CPS focus more on powerful 3C capabilities	The physical world and the cyber world, DTs focus more on virtual models
Cyber–physical mapping[a]	One-to-many correspondence	One-to-one correspondence
Core elements	CPS emphasize sensors and actuator	DTs emphasize models and data
Control	Physical assets or processes affecting cyber representation, and cyber representation controlling physical assets or processes	Physical assets or processes affecting cyber representation, and cyber representation controlling physical assets or processes
Hierarchy	The unit level, system level, and SoS level. A smart production line, shop floor or factory are examples of system-level CPS and DTs; a service platform constitutes SoS-level CPS	The unit level, system level, and SoS level. A complex product can also be considered as a system-level DT; an SoS-level DT covers the product life-cycle
Integration with new IT	Be inseparable from new IT	Be inseparable from new IT. A DT is easier and faster to integrate with new IT compared with CPS

[a] Including two directions—cyber to physical and physical to cyber.

Fast development and evolution of DTw and CPS, fostered by research and technology development, require a more structured approach to the description, analysis and eventually implementation. Id doing that we have to consider not only the technical aspects, but also the operational, human and business implications.

The following model provide a comprehensive representation of an incremental implementation of the CPS approach, specifically in the context of an Industry 4.0 environment, see [CiNe19].
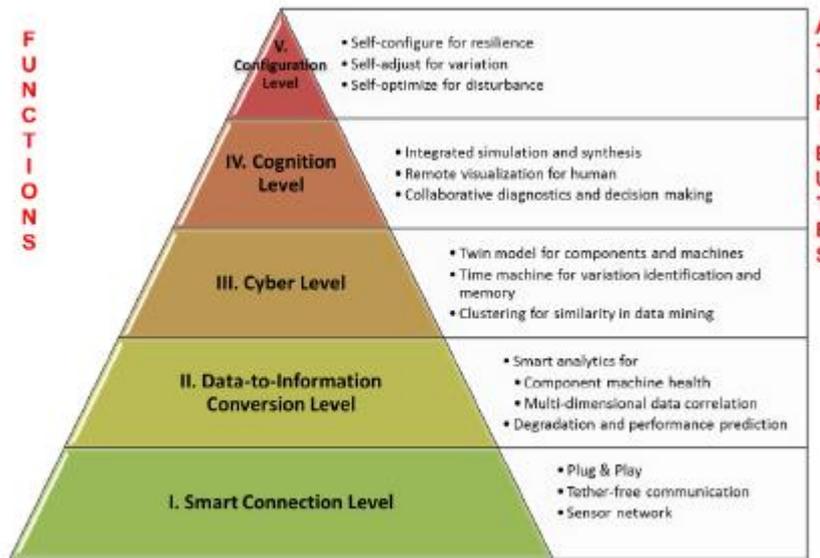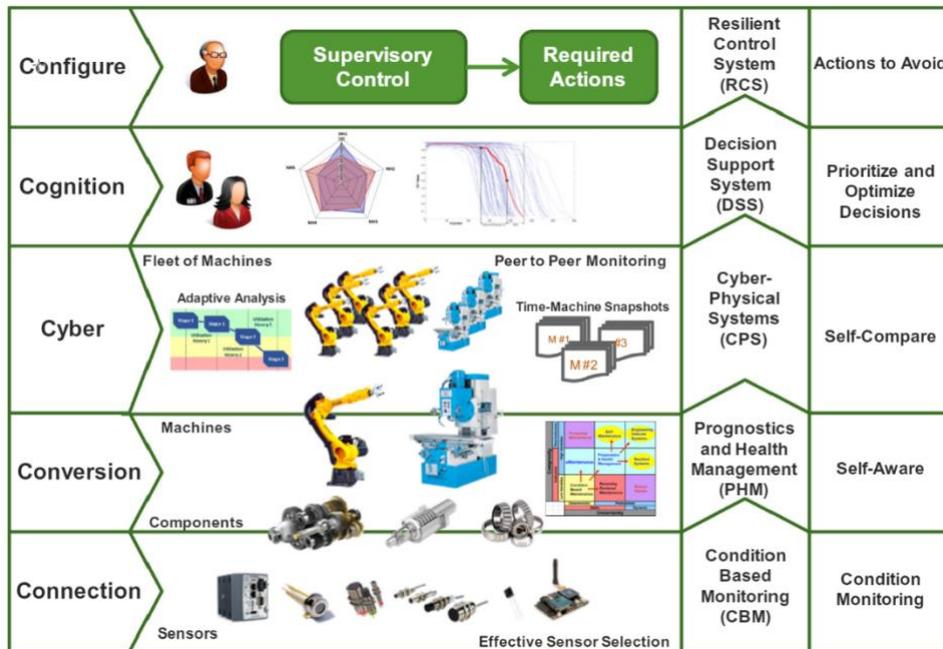

Fig. 1. 5C architecture for implementation of Cyber-Physical System.

Figure 44: 5C Architecture for implementation of Cyber-Physical System, copied from [CiNe19]

For each of the levels it is also possible to identify technological impact as well business and operation impacts, see [CiNe19].

Figure 45: Applications and techniques associated with each level of the SC architecture, from [CiNe19]

It is important to remark how the identified application in order to provide reliable added value services need to satisfy to key attributes, to be connected with a robust, fast and secure way with the field and to adapt the models to the changing situation and configuration in the real world. To such purpose adoption of most advanced technology related to Machine Learning (ML) and generally speaking AI ensure a constant adaptation to changes. At the same way High Performance (HPC) computation capability is needed to execute methods and applications providing the requested services.

Characteristics and requirements for integration of CPS/DTw with a physical environment are summarised below, see [LeBa15]:

1. Ubiquitous connectivity and smart objects: Manufacturing assets should be equipped with smart sensors with the capability of real-time monitoring and data exchange with other elements in the network. These constant data transactions require a secure, reliable, and high-speed platform.

2. Advanced analytics: It is essential to automate the whole process of data pre-processing, perception, analysis, learning, and execution without the need for extensive human interference and manual feature engineering. This process brings self-configure, self-adapt, and self-learning functionalities to the manufacturing systems, which increases productivity, speed, flexibility, and efficiency

3. Cooperative decision making: Data from multiple resources and real-time limitations must be considered to achieve a globally optimal solution. In this process, feasibility, efficiently, and execution plans of different orders are evaluated.

4. Autonomous and rapid model building and updates: Data synchronisation and advanced model mapping between virtual and physical systems guarantee the minimum difference between virtual components and their physical counterparts, which is essential for real-time control, optimisation, forecast, etc.

5. Autonomous disturbance handling and resilience control: Manufacturing systems need to respond to failures autonomously and resiliently in order to prevent catastrophic operational disruptions.

As for the DTw, it is considered to be a new way of managing the industrial IoT. Integrating cloud technologies in DTws holds promise for ensuring the scalability of storage, computation, and communication. BDA, AI, and corresponding algorithms are also seen as important foundations for a DTw. In the exploration of potential DTw applications, new IT and not-IT technologies play a more and more important role, moving from a pure technology perspective towards a holistic approach where many disciplines and skill are required to converge towards a full exploitation of available information. In the following picture it is sketched the DTw/CPS evolution starting for a pure industry related data domain through information elaboration in an IT perspective, but definitively moving towards the broader knowledge domain where not only process/product asset are considered, but also humans are part of the game.
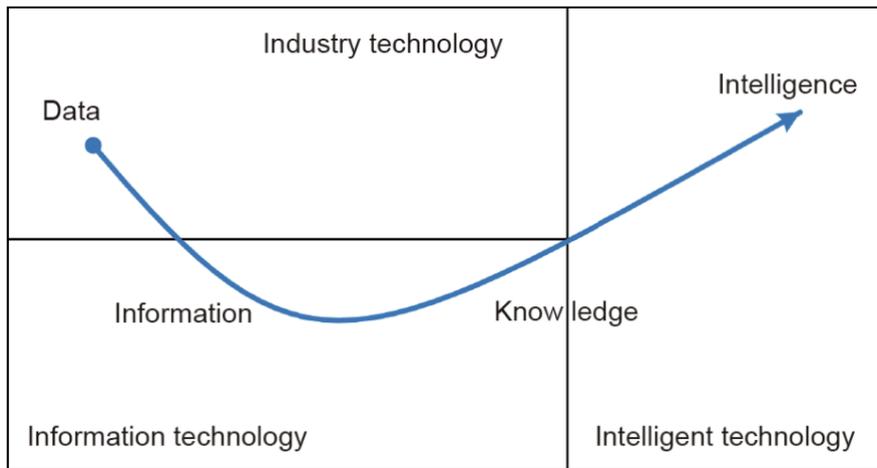
Figure 46: Integration of industrial technology, information technology, and intelligent, copied from [KrKa18]

In such journey 5G technology can play a terrific role, "5G can help support advanced Industry 4.0 strategies by bringing ubiquitous, high speed, reliable, high coverage connectivity to industrial environments and systems " First of all 5G utilizes advanced technologies such as Millimetre Wave and terahertz band, NFV, Wireless Software Defined Network (WSDN), Cloud Radio Access Network (CRAN), and Massive MIMO to provide low latency, high reliability, high transmission rate, high coverage, high security, and scalable networking which can better support the communication demands of future smart manufacturing [LeAz20]. More security mechanism in 5G is addressing some of the concerns for data protection, Frequency Slicing is supporting critical applications requiring specific service level in term of speed and latency, Edge Computing functionality can support distributed computational architecture or Distributed Ledger application. In the following picture a set of functionalities potentially impacted by 5G technology, see [JML20].
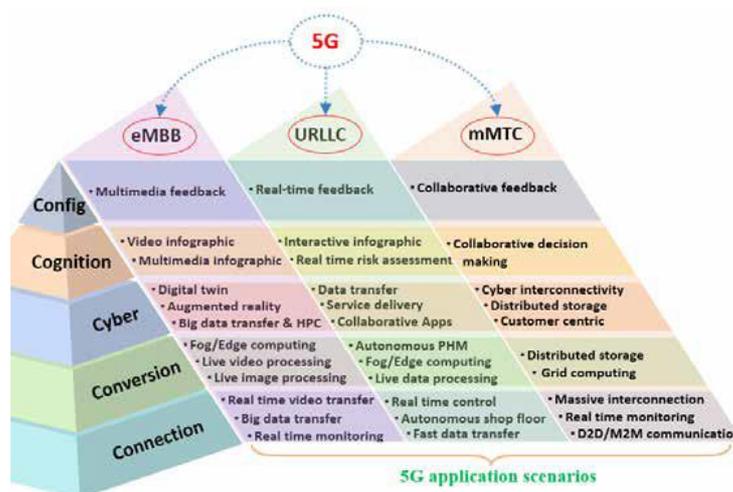

Figure 47: Application Scenarios, copied from [JML20]

3.2 Deep Edge, Terminal and IoT Device Integration in B5G communication infrastructure

This section is related to the Networld2020[8] SNS SRIA [Networld2020-SRIA] and focuses on challenges of the integration of deep edge, terminal and IoT devices in the SNS architecture.

Architecturally, the 'deep edge' with its IoT as well as end user or vertical industry devices is becoming part of the common resource pool, provided as a non-decomposable set of resources by some edge entity, such as an end user, industrial site owner, or a building owner. It is envisioned that tenant-specific resource usage to expand into the deep edge with the same control and data plane considerations and resource management considerations, applying to all those resources. In other words, in principle, we see aspects of controllability of those edge resources to equally apply together with the general programmability for the realization of compute tasks as well as for data and forwarding plane operations through those resources.

However, some edge resources might not directly fit into this vision. For instance, IoT will introduce particular, service-dedicated, possibly intelligent yet resource-constrained components (micro-electronics, battery driven components), which will need a particular consideration for the integration with the rest of the system. Indeed, such IoT components and devices might impose additional requirements on, e.g., volatility and longevity, punctual presence at any moment, persistence, generality, capacities, connectivity, interfaces and APIs from/towards the system. Hence, they might not support direct integration and require particular solutions instead (e.g., gateways or subsystems).

This section focuses on the following objectives, see (section 4.7 of Networld2020 SNS SRIA [Networld2020-SRIA]):

- Future research will need to develop a suitable common model of system-wide representation akin to 'device drivers' in existing computing platforms.

- Future research will need to address edge-specific constraints through suitable scheduling mechanisms that take those constraints into account, while relying on edge-specific control agents enabling the enforcement of the policies underlying the scheduling solutions

- Through research in this space, future solutions to enable an edge resource market that would allow for auctioning the availability of resources to tenants very much like the bidding for white space on a webpage as we know today, basing all interactions on a trusted, auditable, and accountable basis that caters to the dynamics experienced at the edge.

- This will require research into novel programming models and (e.g., policy) languages that not only support all of these services, applications and deployments but also cater to the expected dynamics of the market itself.

- Research is needed for providing new IoT device management techniques that are adapted to the evolving distributed architectures for IoT systems based on an open device management ecosystem.

- In addition, novel programming models and languages are required to support all of these services, applications and deployments. Research challenges in this area include:

  - delivery model and APIs, with effective use of ultra-dense and diverse wired and wireless networks effective management of billions of devices, ensuring they are suitably configured, running appropriate software, kept up-to-date with security updates and patches, and run only properly authenticated and authorized applications.

  - privacy and data management, and the location of processing and data to match legal and moral restrictions on data distribution, access and processing, will be increasingly important.

---

[8] Networld2020 ETP has been renamed to NetwolrdEurope ETP, see: https://www.networldeurope.eu

- policy descriptions, rules and constraints will need to be specified in a form that can be enforced by the infrastructure on the services.

3.3 Edge, Mobile Edge Computing and Processing

This section is related to the Networld2020 SNS SRIA [Networld2020-SRIA] and focuses on Edge, Mobile Edge Computing and Processing challenges.

These approaches require responsive network connectivity to allow "things" and humans to touch, feel, manipulate and control objects in real or virtual environments. Edge processing in the architecture is essential for ultra-low latency and reliability, while the AI processing is transferred at the mobile/IoT device. Research challenges in this area cover open distributed edge computing architectures and implementations for IoT and integrated IoT distributed architectures for IT/OT integration, heterogeneous wireless communication and networking in edge computing for IoT, and orchestration techniques for providing compute resources in separate islands. In addition, built-in end-to-end distributed security, trustworthiness and privacy issues in edge computing for IoT are important, as well as federation and cross-platform service supply for IoT.

In addition, distributed service provisioning will extend also even beyond the edge, i.e., to on-premises devices such as Industrial IoT devices, robots, AGVs, connected cars. Novel forms of dynamic resource discovery, management and orchestration are required, allowing service provisioning to exploit on-premises devices as "on-demand" extensions of resources provided from the core or the edge.

In this framework, novel resource control schemes, balancing between autonomy of devices and the overall optimization and control of the network by the operator(s) will be required, thus innovating the existing collaboration models between different network service providers. This will also allow to take in better account users' context, exploiting the typical co-location of users with on-premises devices and, sometimes, their very tight physical bound. In this sense, this approach will allow designing network services in a more user-centric way.

IoT Distributed and Federated Architectures Integrated with 5G architecture and AI: Further research is needed in novel IoT distributed architectures to address the convergence of (low latency) Tactile Internet, edge processing, AI and distributed security based on ledger or other technologies, and the use of multi-access edge computing. Research challenges include serving the specific architectural requirements for distributed intelligence and context awareness at the edge, integration with network architectures, forming a knowledge-centric network for IoT, cross-layer, serving many applications in a heterogeneous networks (including non-functional aspects such as energy consumption) and adaptation of software defined radio and networking technologies in the IoT.

5G and beyond mobile networks will enable unprecedented density of connected devices many of which will create tremendous amounts of data. As an example, an autonomous car is expected to create data at a rate of estimated 5 terabytes per hour. Transferring these raw data to a central cloud for processing is not feasible for (at least) three reasons:

Bandwidth

If the device is connected via LPWAN (e.g. NB-IoT with an uplink peak data rate of 159 kbit/s[9]) the bandwidth is limited and not suitable to transfer large amount of data (e.g. multimedia data).

Network Congestion

---

[9] See https://en.wikipedia.org/wiki/Narrowband_IoT

With a culminated capacity of the last mile exceeding the capacity of the core network by two orders of magnitude the core is becoming a bottleneck for huge amounts of data to be transferred to the cloud data centres while at the edge there is sufficient capacity available.[10]

Latency

There are applications where latencies beyond the range of hundreds of milliseconds are not acceptable. Multiplayer online gaming is an example which is a driving force in edge development (gamers are paying for latency!). In safety relevant use cases it often is not just a question of "user experience" but a matter of life or death.

Storing (or buffering) raw data locally is often not an alternative either since devices do not have sufficient storage capacity or storage is just too expensive. Taking the example of an autonomous car above and with a current storage price of roughly 20 € per Terabyte to store the raw data of that car would cost 100 € per hour – even without redundancy.

Those restrictions can be overcome by taking content delivery network (CDN) technologies a step further and process data in or near the device by which it is being created (e.g. in a mobile phone or in a surveillance camera). The processing can result in immediate action of an actuator in response to sensor inputs or in condensing data before storing them or sending them to a central cloud. Artificial intelligence comes into play to identify relevant data pattern, but also as a means for network resource optimization and network security. Beyond 5G networks are expected to come with AI already embedded in the network functions[11].

When data are being condensed for transfer or storage this must be done in a manner that potentially valuable information is being retained.

Regulatory requirements may also be relevant for data retention (e.g. in autonomous driving). Such handling of data will be important design decisions when developing edge applications.

Developers are facing competing frameworks to make their apps edge-aware – some of which are provided by large cloud providers (e.g. AWS Greengrass, Azure IoT Edge). To avoid another lock-in, users might consider open-source alternatives like ETSI MEC[12], LF Edge[13], Open Edge Computing[14] or OpenStack[15] (just to name a few).

Developers will also have to deal with different levels of edge computing complexity. One dimension of complexity is the edge-awareness of the application. In the case of edge-unaware applications, developers do not have to deal with the edge specifics and the network is responsible to handle client requests transparently in a manner that those are handled by the server instance with optimum network proximity (just like in today's CDNs). On the other hand, edge-aware applications will have to make use of the available edge-resources by exploiting the specific APIs that are exposed by the edge implementation.

A second dimension of complexity is mobility. When the device is mobile, this is uncritical as long as the edge application is running on the device itself ('device edge'). But if for example the processing is done at the base station ('far edge'), the application context needs to be moved from one base station to another as the user is moving through the mobile network. If roaming between different MNOs comes into play, things even get more complex.

As a side effect, to not send data to a central cloud can be seen as a gain in privacy. However, this presupposes that data security is guaranteed in the edge.

---

[10] See e.g. https://blogs.akamai.com/kr/2018_Edge_Korea_TomLeighton.pdf or https://www.akamai.com/de/de/about/events/edge-highlights.jsp#edgeworld-2019-tom-leighton-through-the-clouds-a-view-from-the-edge (at ~ 13:00 minutes)

[11] See e.g. https://ieeexplore.ieee.org/document/9430853

[12] https://forge.etsi.org/rep/mec

[13] https://www.lfedge.org/

[14] https://www.openedgecomputing.org/

[15] https://www.openstack.org/use-cases/edge-computing/

This, in turn, is not a trivial task, because the attack surface increases enormously and the remote management of the high number of edge devices is a challenge and requires new methods and standards.

Availability can be another benefit of edge computing. Given the edge applications are programmed accordingly they can provide business continuity in situations of loss of network connectivity or downtimes (planned or unplanned) of the cloud data centre.

While edge computing will certainly support the goals of the digital transition, we should not forget about the other side of the medal: sustainability and the green transition. On the positive side of the energy equation, edge computing reduces energy-hungry data transfers. On the downside, the intelligence and processing power required at the edge comes at a (energy) cost. Research should be undertaken on how the net carbon footprint of edge computing could be minimized. When the device is energy constrained (e.g. battery driven) other options like energy harvesting could be taken into consideration.

As the talks and discussions in the workshop *IoT and Edge Computing: Future directions for Europe*[16] have shown edge computing is expected to be the first evolutionary step towards a 'computing continuum' reaching from the cloud data centre to the edge device. Cloud federation as investigated by the European Gaia-X project[17] will allow for flexibility when choosing the cloud vendor preventing vendor lock-ins. Moreover, a split of functions that make up a service will allow to run workloads on the device best suited (e.g. due to the availability of specialized processors like DPUs).

"Edge computing represents the first step towards the decentralisation of Cloud computing, bringing the concept of Federated Cloud to its next evolutionary stage."[18]

As a conclusion, the edge computing paradigm is getting track to deal with some of the shortcomings of the central cloud paradigm. Several technical hurdles need to be overcome with respect to deployment, management and securing of billions of edge devices. Standardisation will be required to avoid islands instead of a continuum. For 5G and beyond mobile networks, edge computing will come in quite naturally to fulfil the promises of ultra-reliability and low latency communications (URLLC) and can be expected to become an integral part of future mobile networks.

3.3.1 Functional Splitting: allowing dynamic computing power allocation for signal processing

The purpose of this section is to provide information on systems oriented to deploy computational power allocation on different parts of the so-called continuum computing. According to Balouek et al[19], this concept aims at "realizing a fluid ecosystem where distributed resources and services are programmatically aggregated on demand to support emerging data-driven application workflows".

Usually, data gathering is made directly for simple parameters coming from direct sensors, but other times the information comes in audio or video format and which made it necessary to allocate some computation power in the nodes, in the Edge or sometimes directly in the Cloud (also computation options in the Fog/Mist can be considered). Another way to focus this problem, as in the node the possibilities to allocate high computation power are few, is to split the signal processing procedure in different blocks and assign (manually or automatically) the computing power for each block (or function) to different parts of the system architecture. This assignment can be managed by an orchestrator, assigning task functions according to the computing resources disposal in the architecture.

---

[16] Workshop of 11 September 2020 hosted by the NGIoT CSA project and organised together with the European Commission and AIOTI, replay and presentations available at https://www.ngiot.eu/event/iot-and-edge-computing-future-directions-for-europe/

[17] https://www.gaia-x.eu/

[18] http://www.pledger-project.eu/FederatedCloud_RA_PP_022021.pdf

[19] D Balouek-Thomert, E. Gibert-Renart, A Reza-Zamani, A Simonet, M Parashar, "Towards a computing continuum: Enabling edge-to-cloud integration for data-driven workflows" Journal of High Performance Computing Applications, Vol. 33(6), pp. 1159-1174, 2019. DOI: 10.1177/1094342019877383

The functional splitting concept is often applied to the 5G network[20], but with this vision, the concept goes beyond the network functional splitting and can be applied to other fields.

In Noriega et al.[21] and Pastor et al.[22], the authors implemented an Edge computing system by using different Raspberry Pi 3 (Rpi3) nodes in order to carry out a performance evaluation with when computing complex audio signal processing metrics directly on Rpi3 nodes, considered as Edge. In Segura et al.[23], authors focus the same problem from the functional splitting perspective with different options in a 5G architecture, see as well the URBAURAMON project.

Other perspectives to face the problem of the improvement of performance in the computation of the complex parameters with a signal processing strategy are: to use a parallel strategy or to use an Artificial Intelligence strategy (e.g. Convolutional Neural Network (CNN)).

In Fayos et al.[24], authors compared a Fog computing system based on different orchestration platforms (i.e. DockerSwarm and Kubernetes) in order to improve performance, for the same complex signal processing problem, with homogeneous and heterogeneous clusters of Small Board Devices. In Salah[25] and El Khafhali et al.[26], the authors focus the efforts in the modelling and provision of the task distribution in the Cloud. In Lopez et al.[27], the authors focused the computing problem by designing a CNN to obtain these parameters and compared its performance with the one of the algorithms in different platforms.

The main challenges associated to the signal processing functional splitting are related to the planned problem and the resources planned in the network (i.e. sampling, windowing, weighting, compression, filtering, etc.). For instance, for audio processing and using ESP32 MCU in the node, we can manage audio sampling, windowing and performing Fourier transform and some other simple operations or functions related to filtering and we can send to the Edge the output information to finish the computing process there. At this point, we need to consider possible delays in the communication but using simple/lightweight protocols (such as MQTT), and using controlled audio/processed chunks, we can obtain affordable delays (i.e. not too high) [5], allowing real-time processing/monitoring. We can also use this procedure for video processing and other temporal related signals but redefining the splitting options to consider the specific problematic of the video processing (e.g. redefining FFT to FFT2D, applying 2D filtering per frame, etc.).

---

[20] D. Harutyunyan and R. Riggio, "Flexible functional split in 5G networks," 2017 13th International Conference on Network and Service Management (CNSM), Tokyo, Japan, 2017, pp. 1-9, doi: 10.23919/CNSM.2017.8255992.

[21] J. E. Noriega-Linares, A. Rodriguez-Mayol, M. Cobos-Serrano, J. Segura-Garcia, F.-C. S., and J. M. Navarro, "A wireless acoustic array system for binaural loudness evaluation in cities," IEEE Sensors Journal, vol. 17, pp. 7043–7052, 2017.

[22] A. Pastor-Aparicio, J. Segura-Garcia, J. Lopez-Ballester, S. Felici-Castell, M. García-Pineda and J. J. Pérez-Solano, "Psychoacoustic Annoyance Implementation With Wireless Acoustic Sensor Networks for Monitoring in Smart Cities," in IEEE Internet of Things Journal, vol. 7, no. 1, pp. 128-136, Jan. 2020, doi: 10.1109/JIOT.2019.2946971.

[23] J. Segura-Garcia, J. M. A. Calero, A. Pastor-Aparicio, R. Marco-Alaez, S. Felici-Castell and Q. Wang, "5G IoT System for Real-Time Psycho-Acoustic Soundscape Monitoring in Smart Cities with Dynamic Computational Offloading to the Edge," in IEEE Internet of Things Journal, doi: 10.1109/JIOT.2021.3063520.

[24] R. Fayos-Jordan, S. Felici-Castell, J. Segura-Garcia, J. LopezBallester, and M. Cobos, "Performance comparison of container orchestration platforms with low cost devices in the fog, assisting internet of things applications," Journal of Network and Computer Applications, vol. 169, p. 102788, 2020. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1084804520302605

[25] K. Salah, "A queueing model to achieve proper elasticity for cloud cluster jobs," in 2013 IEEE Sixth International Conference on Cloud Computing, 2013, pp. 755–761.

[26] S. El Kafhali and K. Salah, "Stochastic modelling and analysis of cloud computing data center," in 2017 20th Conference on Innovations in Clouds, Internet and Networks (ICIN), 2017, pp. 122–126.

[27] J. Lopez-Ballester, A. Pastor-Aparicio, S. Felici-Castell, J. Segura-Garcia, and M. Cobos, "Enabling real-time computation of psychoacoustic parameters in acoustic sensors using convolutional neural networks", " in IEEE Sensors Journal, vol. 20, no. 19, pp. 11429-11438, 1 Oct.1, 2020, doi: 10.1109/JSEN.2020.2995779.
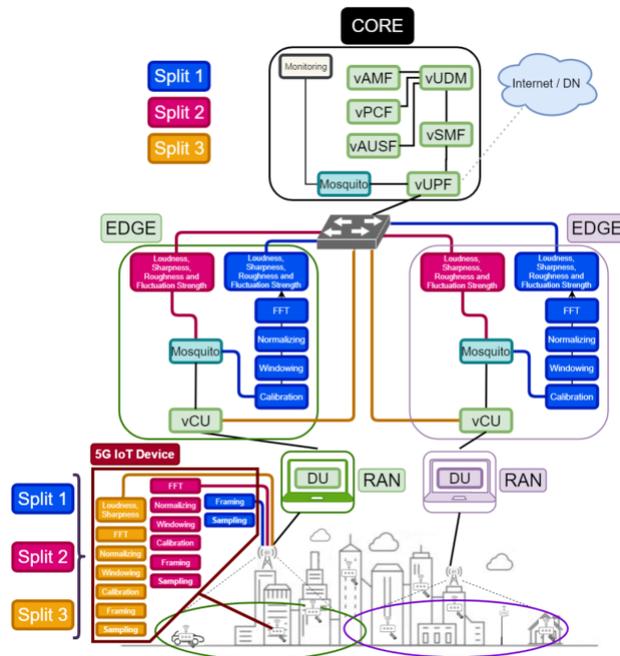
Figure 48: Conceptual diagram of the IoT architecture with different splitting options for the 5G complex metrics calculation system[5]

The 5G IoT infrastructure designed for the soundscape description within the context of a Smart City, considers the following elements or subsystems: a) the node as a 5G IoT sound monitoring device that has connected sensors and collects information, b) the Radio Access Network (RAN) as the radio interface, c) the Edge where some offloading from the device can be applied to allow energy savings and d) the Core where the information is gathered and processed monitoring. **Figure 48** shows a conceptual diagram of these elements with their components, considering the different functional splitting options to compute the metrics for psycho-acoustic soundscape.

The system developed in Balouek et al[1] is an earthquake and tsunami detection and warning global system (by the moment of publication it is deployed in a USA area). Here, the amount of data gathered is huge and the authors propose a ruled-base system for distributing computation loads between Edge and Core and oriented to decentralize the computation, establishing what they call a "virtual slice". This development was made in the context of the GeoSciFramework project (funded by the National Science Foundation).

Another application of this concept is in Rosendo et al[28], where the authors develop a configurable framework for different use cases, but for this project they specify a Smart Surveillance system, achieving very good results in terms of latency and throughput.

In the URBAURAMON project, the main challenges associated to the signal processing functional splitting are related to the planned problem and the resources planned in the network (i.e. sampling, windowing, weighting, compression, filtering, etc.).

In the case of GeoSciFramework project, the proposed architecture is show in **Figure 49** is divided in four layers containing the infrastructure layer (which is divided in two components, such as data producers and computing resources), the federation layer (which defines the relations between the infrastructure components, the streaming layer (which stablishes the rules and constraints for the data processing, indexing and discovery from multiple sources in order to achieve real-time processing, to this end a distributed strategy was followed), and the application layer (which is oriented to manage the data consumers, i.e. applications to deal with data production and delivery – by publication/subscription with MQTT-, establishing the workflow management system and the selection of resources).

---

[28] D. Rosendo, P. Silva, M. Simonin, A. Costan, G. Antoniu. "E2Clab: Exploring the Computing Continuum through Repeatable, Replicable and Reproducible Edge-to-Cloud Experiments". Cluster 2020 - IEEE International Conference on Cluster Computing, Sep 2020, Kobe, Japan. pp.1-11, 10.1109/CLUSTER49012.2020.00028.
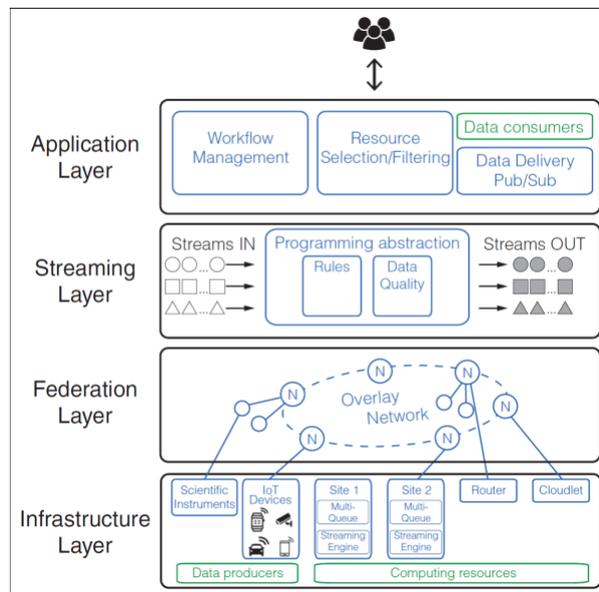
Figure 49: Overall layered architecture of the edge-based data-intensive IoT system.

The E2Clab/OverFlow project, applied an image processing function in a smart surveillance system for counting persons/detecting a specific person or for free parking space detection[29][30] in a Smart City environment.

Also, the use of artificial intelligence in this environment is possible with the distribution of the computing task force in different places of the 5G environment.

3.4   Network and Server security for IoT and edge Computing

This section is related to the Networld2020 SNS SRIA [Networld2020-SRIA] and focuses on Network and Server security for edge and IoT challenges.

The massive deployment of IoT devices and the emergence of 5G technologies in our daily lives are bringing new data-driven and increasingly autonomous scenarios. The realization of these new services requires efficient and effective management of computing and network resources to deal with huge amounts of data and meet the real-time requirements of such applications. To this end, there is a growing trend for the deployment of computing/network resources at the edge of the network, to interconnect the end devices with cloud infrastructures. This results in the cloud-to-edge-to-device spectrum, which represents a *computing continuum*[31] of resources distributed at different network levels.

This trend toward an increasing interconnectivity requires the adoption of automated mechanisms to detect and react against potential cybersecurity attacks. Indeed, in recent years the convergence between Artificial Intelligence (AI) techniques and the adoption of Software-Defined Networking (SDN) techniques is enabling the development of self-protective IoT systems.

To enhance such systems with the ability of detecting potential security attacks or threats, a crucial aspect is the identification of the intended behaviour of each IoT device composing a system.

---

[29] J. Nyambal and R. Klein, "Automated parking space detection using convolutional neural networks," 2017 Pattern Recognition Association of South Africa and Robotics and Mechatronics (PRASA-RobMech), 2017, pp. 1-6, doi: 10.1109/RoboMech.2017.8261114.

[30] G. Amato, F. Carrara, F. Falchi, C. Gennaro and C. Meghini, "Deep learning for decentralized parking lot occupancy detection", Expert Systems with Applications, 72, pp 327-334, 2017. URL: https://github.com/fabiocarrara/deep-parking (Visited on 04/07/2021)

[31] https://ec.europa.eu/digital-single-market/en/news/building-ecosystem-where-iot-edge-and-cloud-converge-towards-computing-continuum

142

Indeed, the use of common machine learning (ML) techniques for the so-called intrusion detection systems (IDS) is based on the definition of the devices' intended or "normal" behaviour to train a certain model (e.g., a neural network). Therefore, the identification of potential actions that are not considered as normal behaviour could be used to infer an attack or threat. In 2019, the Manufacturer Usage Description (MUD)[32] was standardized in the scope of the IETF for the definition of network behaviour profiles for IoT devices. In particular, it describes a data model to restrict the communication from/to a certain device, so that manufacturers are enabled to define the intended network behaviour of their devices. Such behavioural profiles are described by using a set of policies or Access Control Lists (ACL) with the endpoints of the intended communication to reduce the attack surface. Furthermore, the standard specification defines an architecture for obtaining MUD files associated to a certain device containing its intended behaviour. The use of the MUD standard has received a significant interest from Standards Developing Organization (SDO), such as the National Institute of Standards and Technology (NIST), which proposes the MUD standard as a key approach to mitigate denial-of-service (DoS) attacks[33] in home and small-business networks[34].

One of the main potential applications derived from the MUD standard is the development of IDS (Intrusion Detection System) to be considered in IoT scenarios. Indeed, such approach has been considered in recent research activities[35].

In particular, the MUD profiles associated to different IoT devices can be aggregated to build a graph representation of the intended communication in a certain network or system. For example, in a simple approach, graph nodes can be used to represent communication endpoints while edges are used for the interactions between nodes. From the deployment perspective, the use of fog computing could be key to enable an effective detection approach for cybersecurity attacks. Specifically, fog nodes can be used to create a *continuous monitoring* component, so that network traffic of IoT devices can be inspected in real-time. This component could be additionally used to extract the relevant information (i.e., *features*) to be further analysed by an *AI-enabled attack detector*, which is intended to identify potential attacks based on the use of ML techniques. In this context, the use of fog nodes could be used to enable a distributed and cooperative approach for the identification of cybersecurity attacks in IoT-enabled scenarios by performing the tasks associated to network traffic monitoring and attack detection.

Indeed, an important limitation of current approaches to the application of ML techniques for the detection of attacks in IoT, is that they are based on centralized architectures in which a single entity obtains data from the end devices to train a certain model.

This represents a major problem in IoT scenarios, due to the amount and sensitivity of the data that such devices can generate. To address such issue, the use of *federated learning* (FL) is characterized by a collaborative learning process, in which a set of client devices are managed by a central coordinator[36]. However, client devices do not share their data with the coordinator, but only partial updates of the global model that are aggregated by such entity. In each round of training, the coordinator sends information on the current model that is updated by clients through local calculations. This process could foster the compliance of GDPR basic principles. Furthermore, end devices can obtain a more comprehensive overview of the network behaviour since each device obtain information from the other devices in the network.

However, the application of FL in the IoT ecosystem still has to cope with significant challenges related to scalability, heterogeneity and practical aspects, because of the resource constraints

---

[32] E. Lear, D. Romascanu, and R. Droms, "Manufacturer Usage Description Specification (RFC 8520),", 2019

[33] T. Polk, M. Souppaya, and W. C. Barker, "Mitigating IoT-Based Automated Distributed Threats,", 2017

[34] NIST, "Securing Small-Business and Home Internet of Things Devices:NIST SP 1800-15," 2019

[35] S. Singh, A. Atrey, M. L. Sichitiu, and Y. Viniotis, "Clearerthan MUD: Extending Manufacturer Usage Description (MUD)for Securing IoT Systems," inInternet of Things – ICIOT 2019,V. Issarny, B. Palanisamy, and L.-J. Zhang, Eds.Cham: SpringerInternational Publishing, 2019, vol. 11519, pp. 43–57

[36] T. Li, A. K. Sahu, A. Talwalkar, and V. Smith, "Federated learning: Challenges, methods, and future directions, "IEEE Signal Processing Magazine", vol. 37, no. 3, pp. 50–60, 2020.

associated to certain IoT devices[37]. One of the well-known issues of FL is related to the coordinator, which could represent a single point of failure of an FL scenario that could rise the possibility of *poisoning attacks*. Furthermore, poisoning attacks could be also launched by malicious devices by generating false date during the training process. In particular, an attacker could send forged model updates to the coordinator. Therefore, there is a need to ensure only legitimate and authorized devices are enabled to participate in the training process. For this purpose, the use of MUD profiles could be considered, so that only MUD-compliant devices participate during the process[38].

Furthermore, the use of lightweight authentication and identity management schemes for IoT devices is essential to mitigate such attacks. In addition, recent proposals have considered the use of blockchain technology[39], which consists of an immutable transaction and tamper-proof ledger. Thus, instead of sharing the model updates directly with the coordinator, the use of blockchain is proposed to share the global model updates, in order to avoid issues associated to the centralized coordinator entity.

However, the realization and deployment of such ecosystem still needs to be further investigated in the next future to come up with an AI-enabled and automated approach for an effective security attacks detection and mitigation for IoT scenarios.

3.5   Plug and Play Integrated Satellite and Terrestrial Networks

This section is related to the Networld2020 SNS SRIA [Networld2020-SRIA] and focuses on Plug and Play Integrated Satellite and Terrestrial Networks challenges.

Satellite universal coverage, multicasting, and broadcasting capabilities provide enhanced connectivity options and seamless user experience when integrated with the overall 5G system. Satellite systems provide large-scale global connections of services where terrestrial coverage is not available. With an integrated 5G/satellite architecture a truly universal coverage can be achieved [LiGe19]. As IoT density decreases, demands for connectivity change from urban to rural areas, reducing demands on a network, see **Figure 50**.
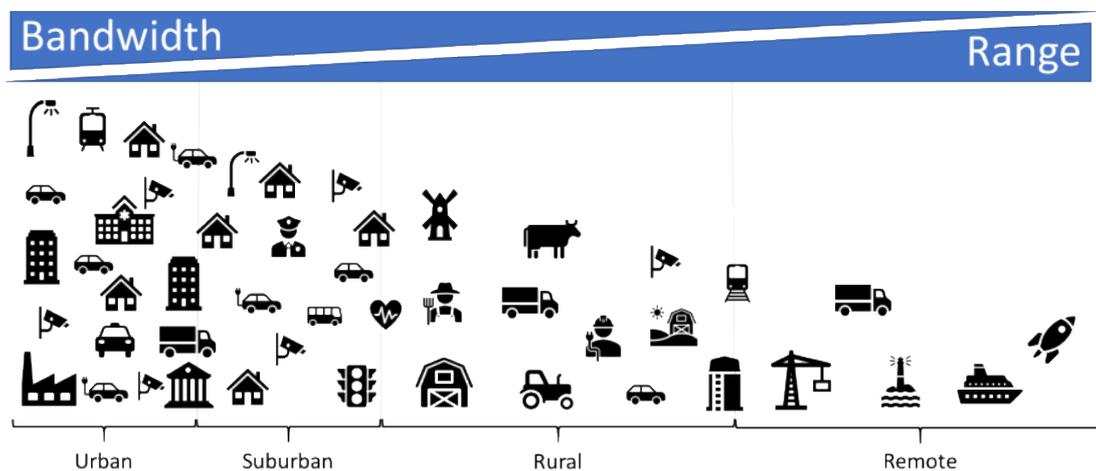
Figure 50: 5G/Satellite Coverage

Traditional Mobile Sat Systems (MSS) like Inmarsat, Thuraya, Iridium, Globalstar have been dominant in the M2M/IoT market, using their L-band spectrum with a focus on mobile and maritime applications. In the last 10 years they realised 3.5 - 4 million satellite IoT terminals in the

---

[37] Imteaj, A., Thakker, U., Wang, S., Li, J., & Amini, M. H. (2020). Federated learning for resource-constrained iot devices: Panoramas and state-of-the-art. arXiv preprint arXiv:2002.10610.

[38] Feraudo, A., Yadav, P., Safronov, V., Popescu, D. A., Mortier, R., Wang, S., ... & Crowcroft, J. (2020, April). CoLearn: Enabling federated learning in MUD-compliant IoT edge networks. In Proceedings of the Third ACM International Workshop on Edge Systems, Analytics and Networking (pp. 25-30).

[39] Lu, Y., Huang, X., Dai, Y., Maharjan, S., & Zhang, Y. (2019). Blockchain and federated learning for privacy-preserved data sharing in industrial IoT. IEEE Transactions on Industrial Informatics, 16(6), 4177-4186.

field. With the availability of Ku-band and Ka-band satellite connections provides higher through-put to meet the demand on of the IoT sector such as fixed satellite systems like Eutelsat, Intelsat or Asiasat. Their higher bandwidths provide backhaul services connecting terrestrial local area IoT networks (e.g., NB-IoT, Lora, WiFi, BT) from high density sensor networks to the internet, see [Satell-market].

New satellite players take advantage of the new cubesat technology (using a range of UHF, VHF, S-band, and Ku-band services) to bring down their service costs, while the Low Earth Orbit allows the use of low power modems to connect the ground sensors, see [KoLa20].

Nanosatellites are defined as any satellite weighting less than 10 kilograms. They all are based on the standard CubeSat unit, namely a cube-shaped structure measuring 10x10x10 cm with a mass of somewhere between 1 kg and 1.33 kg. This unit is known as 1U. As the number Internet of Things (IoT) devices and Machine-to-Machine (M2M) communications increases at an exponential right rate. No communications system can provide end to end connectivity and satellite systems create the opportunity to provide extended coverage, see [NASA-cubesats].

Companies such as Astrocast, Myrioata, Lacuna, Kineis, Kepler Communications, Swarm technologies and Hiber provide service features, low cost, low power, low latency, makes them well suited for Direct-To-Satellite services.

For satellite systems to integrate with 5G networks the architecture will need to address a number of specific issues namely, see e.g., [ISTINCT]:

- Diversification of the spectrum usage across multiple technologies
- Edge networks to reduce the impact of the backhaul in the end-to-end system
- Adapted data path protocols to massive communication environments
- Application protocols adaptation through the virtualization environment
- Addressing the M2M communication needs in an efficient manner
- Participation within the main standardization organizations: 3GPP, ETSI NFV,ETSI MEC, IETF, ONF

3.5.1 Satellite connectivity for global IoT coverage

Today, there are 1.7 billion cellular IoT devices active worldwide. By 2026, there will be 5.9 billion according to Ericsson [Ericsson20], an increase of nearly 350%. Given this tremendous growth, it is clear that the ability to connect diverse IoT device types, with different needs, at massive scale and with global coverage, is urgently needed.

Mobile network coverage is mostly focused on areas with mid to high population density. Areas with low density of population are underserved because of the small or null return on investment required to cover such regions. Currently only 30% of the Earth's landmass, or 10% of the Earth surface has mobile network coverage.

IoT applications such as vehicle monitoring, asset tracking, agricultural sensors and infrastructure monitoring cannot be deployed or used where there is no terrestrial network. Therefore, benefits provided by IoT applications cannot currently be achieved in large portions of the Earth surface.

The capability of satellites to provide global coverage makes them an excellent choice to address the lack of coverage in low populated, isolated and remote areas. The combination of satellite communications together with 3GPP standards offer the possibility to integrate terrestrial and non-terrestrial networks in an easy and simple way.

There are already satellites today that offer global connectivity services for IoT but the communication protocols used are not standard, which requires the development of dedicated terminals, and are typically dedicated to specific vertical solutions. Also, current satellite solutions do not integrate with existing IoT terrestrial networks and, finally, its cost does not meet the price points required for massive IoT deployment.

The market today is demanding standard solutions based on roaming, such as 5G, which are interoperable with terrestrial networks, avoid vendor chipset and service provider lock-in, benefit from massive scale deployment and chipset manufacturers diversity. These requirements

provide the lowest cost solution on chipset and service costs, reduce dependencies on manufacturers and service providers and protect investments on sensors. Combining terrestrial and satellite networks under 5G makes it possible to ensure seamless connectivity using the best available network at any time, see **Figure 51**.
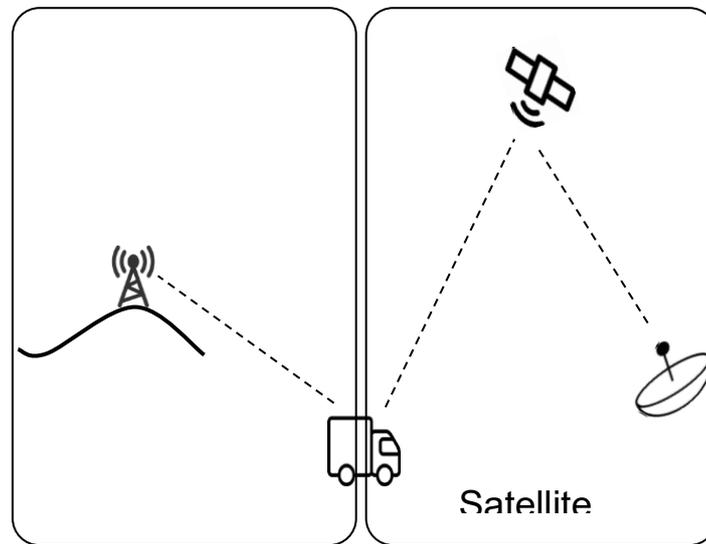
Figure 51: Integrated terrestrial and satellite IoT networks

3.5.2   Evolution to 5G IoT over satellite

While traditionally satellite and terrestrial standardization have been separate processes from each other, the satellite communications industry is nowadays strongly involved in the 5G standardization process led by 3GPP in a quest towards achieving a higher layer operational integration and high degree of radio interface commonality between non-terrestrial networks (NTN) and 5G radio access technologies. Studies on satellite access began in 3GPP a few years ago in the context of Rel. 14 and Rel. 17, to be finalized by mid-2022, will be the first version to support 3GPP standards running over non-terrestrial networks. Specifically, Rel-17 is expected to come with an adaptation of the 5G New Radio (NR) protocol for NTN (this work is already at normative phase, after completion of the study phase) as well as adaptation of the NB-IoT and eMTC protocols for NTN (this work is at study phase).
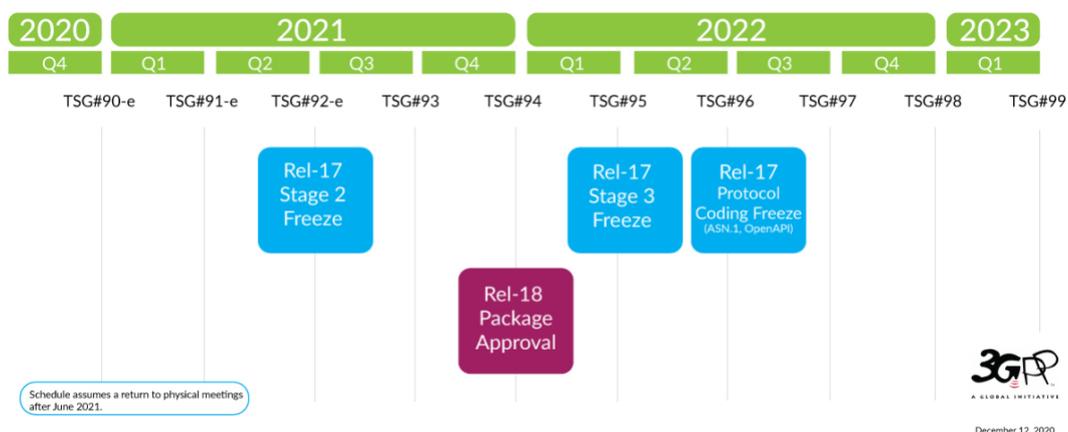
Figure 52: 3GPP Release 17 timeline, copied from 3GPP

Today it is not clear whether Release 17 study phase of IoT over NTN will be moved to normative phase in 3GPP RAN plenary meeting TSG#92-e that will take place in June 2021 for a deployment timeframe for Rel-17 and IoT services over satellite around 2023-24.

146

The following opportunities to provide input on the 3GPP SA1 group, focusing on services, will be in S1-94 in May/July and S1-95 in August, which will address services for Release 18, see **Figure 52.**

3.5.3   IoT devices

3GPP current study items plan that IoT devices will support both terrestrial and non-terrestrial networks on the same device for integrated and seamless connectivity. This makes it possible for the device to select the best and most cost-effective network at any given time. By having a single chipset capable of connecting to mobile and satellite networks it is not necessary to implement two different RF chains that increase complexity and cost of the IoT device. Moreover, the chipset can benefit from the economies of scale provided by all mobile and satellite IoT devices using the same chipset. Typically, the terrestrial network will be used when there is coverage and the device will roam into the satellite network when there is no mobile terrestrial network available.

Testing performed by Mediatek and Inmarsat in August 2020 [3GPP-TSG-RAN89E] show that IoT Satellite communication could be possible with current NB-IoT chipsets. If this is confirmed then existing IoT devices using NB-IoT could use satellite connectivity without having to modify or replace its current hardware just with a firmware update. The firmware update would support the waveform required to cope with the impairments of the satellite connection providing backward compatility, while switching from one network to the other will be supported by already existing 3GPP roaming support.

Satellites providing 5G IoT connectivity may use transparent or regenerative payloads in the satellite. LEO satellites will tend to use regenerative payloads because of the discontinuous connectivity to the core and the needs of 5G to establish connections with the terminals/IoT devices. GEO satellites can use either transparent or regenerative payload on the satellites as they have the possibility to connect to a base station on the ground for signalling.

3.5.4   IoT communication satellites

Traditionally satellite communications have been delivered by Geostationary satellites. Advances in space technology have opened the possibilities for LEO, Low Earth Orbit, satellites to also provide communication services. For this reason, there will be several options for IoT satellite services and its selection will depend on the requirements of the IoT application such as bandwidth, delay tolerance and service continuity.

In contrast with services designed to provide high data rates and continuous service, which are likely to require dense constellations (e.g. in the order of hundreds or more) of high-capacity satellites, NB-IoT solutions with sparse LEO constellations (e.g. in the order of tens of satellites) of CubeSats or similar platforms are anticipated to be a compelling approach to address the needs of many IoT and M2M applications. In particular, there is a wide range of delay-tolerant IoT/M2M applications that do not require continuous service coverage and that generate short, infrequent messages that can be properly addressed with such solutions. For example, in smart agriculture applications, small messages, few messages per day, large delays are not a service problem and can be perfectly achieved by a satellite network not offering continuous coverage. More examples are maritime use cases for non-critical asset tracking where today a data logger is already used, livestock monitoring during pasture in rural areas, and in general any non-critical asset tracking, environmental monitoring and infrastructure monitoring.

Satellite constellations based on CubeSat technology can benefit from low complexity and cost-effective solutions to offer the IoT services, and its required infrastructure, being discussed in this report. Together with the increase of launch opportunities due to new launchers being available and its reusability, this new model, sometimes referred as the New Space model, has greatly increased the number of satellites being built, launched and deployed.

With the increased number of satellites and satellite constellations being deployed at the moment, it is imperative that the satellite design includes its deorbit once its mission has finished in order to minimize the space debris. Satellites must follow ISO 24113:2019 Space Systems-Space

Debris Mitigation Requirements [ISO 2413] and ESA Space Debris Mitigation Compliance Verification Guidelines ESSB-HB–U-002 [ESA ESSB HB –U 002].

3.6   Autonomous and Hyper-connected On-demand Urban Transportation

The transportation domain is ongoing an evolution towards increasing levels of connectivity and automatism. This is the so-called Collaborative, Connected and Automate Mobility (CCAM) paradigm[40]. In this evolution, vehicles will be increasingly connected through different wireless standards like ITS G5 and LTE-V2X but they will also benefit by increasing level of automatism[41]. While the possibility of having fully automated vehicles (level 5 of the J3016 standard)[42] may still take considerable time to happen, levels 2, 3 and 4 are more near deployment in the market or they are already deployed in the market[43]. There are considerable expectations for these new technologies and many studies and reports have identified a number of key benefits for the deployment of these technologies from the obvious and primary benefit to improve the safety conditions in the road to improvement in traffic management, improve compliance to regulation and so on.

The connectivity trend and the automated vehicle trend have evolved from different origins as the first (connectivity) trend is focused on providing connectivity to the vehicle for a variety of applications including safety while the second (automated vehicle) trend is focused on applying artificial intelligence to the processing and analysis of the data originating from the sensors to improve the awareness of the vehicle intelligence.

There is a logical link between the two trends because the connectivity technologies can provide useful information to the automated vehicles for different levels of automation, so that it is an additional input to the artificial intelligence component in the vehicle[44].

There are two main connectivity technologies: short range communications which provides fast communication between vehicles (V2V) and vehicles to infrastructure (V2I) and long-range communication (e.g., 3GPP) where the vehicle can be both the source of information to back-end offices for various applications (e.g., traffic management) but it can also be a recipient of information (e.g., weather conditions). V2X has been traditionally designed using the 802.11p standard[45] while long range communication can be provided by cellular networks. On the other side, there are ongoing discussions on the possibility that 3GPP can also be used for V2X using Device 2 Device (D2D) protocols.

For example, in USA, 3GPP has also been proposed for V2X communication leading to a possible coexistence of the two technologies at least in some geopolitical areas (e.g., USA)[46]. Additional details on the debate on ETSI ITS G5 versus 3GPP LTE-V2X can also be found in section 3.2 of the AIOTI report "IoT Relation and Impact on 5G"[47]. The security (authentication and integrity) of V2X has been designed and described in ETSI and IEEE standards[48] and they may rely on a Public Key Infrastructure (PKI).

The security of cellular networks for long range communication can be based on the authentication, integrity and encryption already described in the 3GPP standards even if it was designed for a different use case.

---

[40] Alonso Raposo, M., Grosso, M., Després, J., Fernández Macías, E., Galassi, C., Krasenbrink, A., ... & Ciuffo, B. (2018). An analysis of possible socio-economic effects of a Cooperative, Connected and Automated Mobility (CCAM) in Europe. European Union.

[41] Weber, R., Misener, J., & Park, V. (2019, May). C-V2X-A Communication Technology for Cooperative, Connected and Automated Mobility. In Mobile Communication-Technologies and Applications; 24. ITG-Symposium (pp. 1-6). VDE.

[42] SAE, S. (2014). J3016 standard: taxonomy and definitions for terms related to on-road motor vehicle automated driving systems.

[43] Yang, CY David, Kaan Ozbay, and Xuegang Ban. "Developments in connected and automated vehicles." (2017): 251-254.

[44] Tong, W., Hussain, A., Bo, W. X., & Maharjan, S. (2019). Artificial intelligence for vehicle-to-everything: A survey. IEEE Access, 7, 10823-10843.

[45] Jiang, D., & Delgrossi, L. (2008, May). IEEE 802.11 p: Towards an international standard for wireless access in vehicular environments. In VTC Spring 2008-IEEE Vehicular Technology Conference (pp. 2036-2040). IEEE.

[46] Bey, T., & Tewolde, G. (2019, January). Evaluation of DSRC and LTE for V2X. In 2019 IEEE 9th Annual Computing and Communication Workshop and Conference (CCWC) (pp. 1032-1035). IEEE.

[47] AIOTI Report. IoT Relation and Impact on 5G. Release 3.0. https://aioti.eu/wp-content/uploads/2020/05/AIOTI-IoT-relation-and-impact-on-5G-R3-Published.pdf

[48] Fernandes, Bruno, João Rufino, Muhammad Alam, and Joaquim Ferreira. "Implementation and analysis of IEEE and ETSI security standards for vehicular communications." Mobile Networks and Applications 23, no. 3 (2018): 469-478.

Automation technologies include the artificial intelligence component, which is used both for a) data analysis of the data originating from the sensor (e.g., camera, LIDAR, inertial measurement units) and b) composing the awareness context of the vehicle and c) taking a decision on the action to take (e.g., avoid a pedestrian).

Beyond the technologies underlying these trends, we also investigate here the potential impacts (e.g., societal) and the potential applications of the combined connectivity and automated concepts, otherwise called CCAM (Cooperative, connected and automated mobility).

At the highest level of automation (level 5 in J3016), the concept of vehicles sharing have been proposed by various sources. In this concept, the vehicle is not owned and driven (for automation levels below 5) by a single proprietary but it can be shared among different users, thus leading to a new economy model where ownership is replaced by pay-by-use.

The emergency of such sharing models can be applied not only to passenger's vehicles but also to commercial vehicles and to public transportation where the vehicles will be owned by the government. Such sharing models poses new challenges not only because they can be economically disruptive (businesses may disappear) but they can also generate great risks from a privacy and security point of view. From a privacy point of view, it is imperative that the data on the passengers is not disclosed or accessible to un-authorized party. From a security point of view, it is necessary that shared automated vehicles cannot be compromised and used for criminal activities[49].

The recent terrorist attacks where commercial vehicle was used to kill pedestrians[50] could be replicated with a shared vehicle driven remotely or with a driving plan inserted in the automated vehicle driving engine by a terrorist or a criminal.

Then, for these reasons or other reasons, it is possible that shared vehicles will be submitted to stringent type approval processes even more than conventional vehicles. The integration of shared commercial vehicles with other means of transportation would also improve the efficiency of the supply chain as the so called "last mile" delivery can be automated through this concept.

Apart from the driverless vehicles (i.e., level 5) the lowest levels of automation can still generate new applications which would greatly benefit the road transportation sector. We can identify just few of them. The presence of sensors in the vehicle and artificial intelligence components can be used to support more sophisticated applications of traffic management where the data from sensors is conveyed to back-end traffic management applications where the traffic conditions (e.g., traffic signs, urban public transport) can be made more efficient on the basis of the real-time received data. In addition, vehicles equipped with inertial measurement units can provide real-time information on the conditions of the road surface for road maintenance purpose or to improve safety (e.g., slippery conditions due to rain can be analysed and communicated to other vehicles in the region). In another example, the findings from the artificial intelligence components of the vehicle (e.g., optimal weights of the deep learning algorithms) can be shared among the AI component of the vehicles to improve driving efficiency.

---

[49] De La Torre, G., Rad, P., & Choo, K. K. R. (2020). Driverless vehicle security: Challenges and future research opportunities. Future Generation Computer Systems, 108, 1092-1111.

[50] https://www.history.com/this-day-in-history/2016-nice-terrorist-attacks

For example, the poor lighting or surface conditions in a specific urban area can be mitigated by making the Artificial Intelligence (AI) components of different vehicles travelling in the area to share the model parameters through federated learning[51]. As in other contexts, it is important that the integrity of the exchanged data is protected because false data can compromise the functioning of the AI components and therefore the safety of passengers and pedestrians.

Finally, we would like to highlight that the emergency of CCAM would require complex data management and analysis systems and infrastructures as the amount of data originating from the vehicles can be massive. We also note that the tracking of the history of the vehicles is particularly important for maintenance purposes or for compliance to regulations because of the long lifetime of the vehicles. Then, technologies like the Blockchain with its properties of decentralization, transparency, and immutability can be quite beneficial in this context[52].

3.7   Opportunities for IoT Components and Devices

This section is related to the Networld2020 SNS SRIA [Networld2020-SRIA] and focuses on Opportunities for IoT Components and Devices challenges.

Deploying and managing a large set of distributed devices with constrained capabilities is a complex task. Moreover, updating and maintaining devices deployed in the field is critical to keep the functionality and the security of the IoT systems. To achieve the full functionality expected of an IoT system, research should be done in advanced network reorganization and dynamic function reassignment. Research is needed for providing new IoT device management techniques that are adapted to the evolving distributed architectures for IoT systems based on an open device management ecosystem in a high threat landscape.

Components (micro-electronic components) and devices mainly for IoT and vertical sector applications are essential elements of future secure and trusted networks and to support the digital autonomy of Europe.

With respect to the increasing demand and expectation of secure and trusted networks, especially for critical infrastructures, there should be European providers for such devices as an additional source to latest technologies to complement the European value chain and mitigate the existing gaps.

3.7.1   Approach for components

European semiconductor players are stronger in IoT and secured solutions, while mass-market oriented market are dominated by US or Asian players. For European industry to capture new business opportunities associated with our connected world, it is crucial to support European technological leadership in connectivity supporting digitisation based on IoT and Systems of Systems technologies.

Increasingly, software applications will run as services on distributed systems of systems involving networks with a diversity of resource restrictions.

It is important to create the conditions to enable the ecosystem required to develop an innovative connectivity system leveraging both heterogeneous integration schemes (such as servers, edge device) and derivative semiconductor processes already available in Europe.

Smart services, enabled by smart devices themselves enabled by components introducing an increasing level of "smartness", will be used in a variety of application fields, being more user-friendly, interacting with each other as well as with the outside world and being reliable, robust and secure, miniaturised, networked, predictive, able to learn and often autonomous. They will be integrated with existing equipment and infrastructure - often by retrofit.

---

[51] Chai, H., Leng, S., Chen, Y., & Zhang, K. (2020). A hierarchical blockchain-enabled federated learning algorithm for knowledge sharing in internet of vehicles. IEEE Transactions on Intelligent Transportation Systems.

[52] Baldini, G., Hernández-Ramos, J. L., Steri, G., Neisse, R., & Fovino, I. N. (2020). A Review on the Application of Distributed Ledgers in the Evolution of Road Transport. IEEE Internet Computing, 24(6), 27-36.

Enabling factors will be: Interoperability with existing systems, self- and re-configurability, scalability, ease of deployment, security, sustainability, and reliability, will be customised to the application scenario.

Related to technological game changers in 5G network infrastructure, Europe strengths are RF SOI and BICMOS technologies for cost-effective GaAs replacement, FD-SOI for integrated mixed signal System on Chip.

The 5G technologies and beyond utilise the sub-6 GHz band and the spectrum above 24 GHz heading to millimetre-wave technology moving towards 300 GHz and Terahertz frequencies for 6G technologies.

The design of electronic components and systems to provide the 5G and beyond connectivity have to take into account the new semiconductor processes for high-speed, high-efficiency compound semiconductor devices considering the significant increases in the density of wireless base stations, wireless backhaul at millimetre wave frequencies, increased transport data rates on wired networks, millimetre wave radios in 5G equipment and multi-frequency/multi-protocol IoT intelligent nodes to support higher data rates, more devices on the network, steerable beams resulting from massive MIMO antennas, low power consumption and high energy efficiency.

It is expected that the mobile and intelligent IoT devices to provide edge computing capabilities and intelligent connectivity using multi-frequency/multi-protocol communications technologies. Cellular IoT devices covering higher frequencies need to integrate microwave and analogue front-end technology and millimetre wave monolithic integrated circuits (MMIC).

The development of 5G technologies and beyond requires semiconductor technologies that are used for RF devices, base stations, pico-cells, power amplifiers to cover the full range of frequencies required. Horizon Europe partnerships Smart Network and Services JU (SNS JU) and Chips JU have to address the development of III-V semiconductors-based GaAs, GaN, InGaAs, SiC semiconductor technologies to implement new components, devices and systems to have the edge in efficiency and power usage needed for base stations.

The new devices for 5G technologies and beyond need to combine RF, low operating power, thermally and energy-efficient, small form factor and heterogeneous integration of different functions. These new requirements push for creating new components based on multi-chip modules and Silicon in Package (SiP) and various technologies that combine the capabilities of silicon CMOS with III-V semiconductors.

The focus for new 5G and beyond connectivity IoT devices is on providing new components including hybrid electronic circuits able to operate with better stability, less noise, providing increase functionality, complexity, and performance. The new functionalities include stronger security mechanisms and algorithms integrated into the devices and components and designed for easy implementation of end-to-end security at the application level.

Activities need to be aligned with the Chips JU to develop 150 mm and beyond wafers for III-V semiconductors on Silicon to provide the components for 5G and beyond wireless cellular networks and devices for providing optimum use of available bandwidth for millimetre-wave and higher frequencies.

Components must be designed to meet the security requirements of critical infrastructure as required on high level by the NIS directive[53], Cybersecurity Resilience Act[54] and the US Executive Order on Improving the Nation's Cybersecurity[55].

---

[53] NIS Directive, https://digital-strategy.ec.europa.eu/en/policies/nis-directive
[54] Cybersecurity Resilience Act, https://www.chips-ju.europa.eu/
[55] Executive Order on Improving the Nation's Cybersecurity, https://www.whitehouse.gov/briefing-room/presidential-actions/2021/05/12/executive-order-on-improving-the-nations-cybersecurity/

ENISA has published several best practices documents on IoT security and securing the IoT supply chain[565758], as well as other organizations such as NIST[5960] and GSMA[61] Specific to 5G networks the EU Cybersecurity Act will mandate certifications for specific components in 5G networks[62], particularly on the network level but users of 5G IoT networks are expected to require string security functions to enable the vertical applications. For components this means that they must include technology enabling high security such as cryptographic hardware, secure updates and a secure component supply chain from cradle to grave. There is an opportunity in being able to early on supply the security needed by future networks and applications.

SNS JU will not directly be involved in component research, development and design. However, the research and development in the SNS JU will enable other initiatives to provide the know-how and later the design and production of communication and computing components.

These activities will help to facilitate the re-launch of the micro-electronics industry in the ICT domain in Europe by means of cooperation with the Chips JU by promoting the development of European added value embedded solutions for innovative and secure applications. SNS JUs will develop the communication know-how and IPRs and will provide algorithms to the micro-electronics industry, which will be dealing with the design and production. With this approach ongoing activities in the Chips JU can be leveraged. From the SNS JUs perspective that could be a fabless approach. A joint effort of different Partnerships under Horizon Europe will involve the appropriate expertise from different communities.

3.7.2 Approach for devices

Devices and especially end devices for IoT and vertical applications including critical infrastructures are an essential part of future networks. In addition to components, they also must fulfil a high security level.

The SNS JU is enable and validate, among others, specialised devices for IoT and sensor systems especially for vertical sectors by leveraging system on chip activities and specifying the way they communicate in the network/systems as well as controlling them and integrating them in their operational systems in vertical (and as well cross- vertical) application domains by means of cooperation with the Chips JU and leveraging AIOTI activities.

System on chip activities can be leveraged for such industrial device activities. The close cooperation between vertical sectors and the ICT industry in Europe will support the development of entire communication and networking solutions in Europe. These activities offer opportunities for start-ups to design communication modem chips and other components devised for many vertical applications.

Devices must be designed with a security first approach, considering the whole life cycle of devices. Especially for critical infrastructure this will be mandated early on but these requirements will also affect other devices as the threat landscape continues to evolve, expanding on the opportunity. For devices this means that manufacturers must adopt a holistic view on supply chain security including all components that go into the device. The device must contain enough security functionality to enable the user to adopt zero trust and zero touch architectures and paradigms including verifying the supply chain, secure deployment of devices and secure life cycle management of devices over the whole device lifecycle, including potential ability to upgrade to future post quantum cryptographic algorithms.

---

[56] ENISA Guidelines for Securing the Internet of Things, https://www.enisa.europa.eu/publications/guidelines-for-securing-the-internet-of-things
[57] ENISA Good Practices for Security of IoT, https://www.enisa.europa.eu/publications/good-practices-for-security-of-iot-1
[58] ENISA Baseline Security Recommendations for IoT, https://www.enisa.europa.eu/publications/baseline-security-recommendations-for-iot
[59] NISTIR 8259A: IoT Device Cybersecurity Capability Core Baseline, https://csrc.nist.gov/publications/detail/nistir/8259a/final
[60] NISTIR 8259: Foundational Cybersecurity Activities for IoT Device Manufacturers, https://csrc.nist.gov/publications/detail/nistir/8259/final
[61] GSMA IoT Security Guidelines and Assessment, https://www.gsma.com/iot/iot-security/iot-security-guidelines/
[62] Securing EU's Vision on 5G: Cybersecurity Certification, https://www.enisa.europa.eu/news/enisa-news/securing_eu_vision_on_5g_cybersecurity_certification

3.7.3   Requirements for IoT devices

Devices with IoT gateway capabilities in support of different IoT connectivity modes, both at local and public network level. In particular for each supported vertical industrial domain and as well cross vertical industry domains:

- requirements will be derived on which software and hardware capabilities and characteristics these multi-modal IoT devices and network elements should support, when integrated and used into the 5G and beyond 5G network infrastructures. Considering that these IoT devices support e.g., wireless technologies that are non-5G and beyond 5G radio technologies, such as Bluetooth, Wi-Fi, ZigBee, LoRa, Sigfox.

- integration and evaluation activities of these multi-modal IoT devices and network elements in the 5G and beyond 5G network infrastructures will be planned and executed.

- Hardware requirements for IoT Devices:

  - Requirements applied for each supported vertical industry domain and as well cross vertical industry domains when integrated and used into the 5G and beyond 5G network infrastructures.

  - At least three different frequency bands for sub-1 GHz, (700 MHz), 1 - 6 GHz (3.4 - 3.8 GHz), and millimetre-wave (above 24 GHz) and integrate multiple protocols in addition to cellular ones.

  - Functional and non-functional requirements, such as high data capacity, highest levels of reliability (connectivity), fast reaction times (low latency), sensing/actuating, processing and storage capabilities; low power consumption.

  - Strong security functionality with hardware cryptographic security modules, initial device identities and upgradable cryptographic algorithms.

3.8   EU legislative framework

Many of the gaps identified for the coverage of remote areas, or with very little population density are still not properly addressed today, where no public network coverage is available. This requires the need to create new technological solutions, where you can combine resources from different suppliers. One of the options could be linked to the use of equipment in the fields, which could be used as relays to reach an area covered by a tower. However, the implementation of such solutions should not modify the behaviour of the integrity of such equipment.

Many conformity assessments for safety and security are today supported by the Original Equipment Manufacturer (OEM) to validate the compliance of an equipment to get the CE marking and homologations or certifications. These requirements are applied on equipment used in the fields and/or potentially used on a public network.

The use of the European and international standards is needed to allow proper risk assessments under the future regulation for machineries replacing the current Machinery Directive (2006/42/EC). Integrating new technologies (IoT devices, AI/ML, cyber-security, autonomous features, etc.) into the Essential Health and Safety Requirements, while maintaining high levels of safety and security, and protecting the OEM against potential litigations, is challenging. This comes to the proposal of a valid business case to engage OEM in standard developments with a good legislation. The ultimate goal is to protect the end user while mitigating the risk of misuse of the equipment.

With the connectivity of such equipment, the OEM sometimes can hardly differentiate which legislation is on top of the other, when he reviews the Radio Equipment Directive, the Electro-Magnetic Compatibility directive, and the Machinery Directive. This is the reason why the technical specifications to implement such relays will determine a hierarchy and include the compliance to these European legislations to address these risks at the same time.

Part of these requirements includes privacy and trust in the data transferred. The data governance is not part of the scope and the solution to develop is to provide the access to an area covered by a telco provider through the relays supported by the equipment in the fields.

The requirements mandated in the new EU legislation such as Cybersecurity Resilience Act and AI Act[63] also need to be taken into account.

3.9   Key AI concepts

Key AI concepts that play a significant role in the development of AI are:

Large Learning Models (LLMs): such as GPT-4 [AcAd23], PaLM [ChNa23] DeepSeek and Kimi K2 are trained on massive datasets of text from books, web content, and other types of contents. These types of models have large capabilities related to natural language understanding, answering questions, summaries and summarization, dialogue logic, and in some cases as well symbolic reasoning [RoTs25a]. Within the AI Agent architectures, LLMs serve as the primary decision-making engine, allowing the agent to parse user queries, plan multi-step solutions, and generate human-like responses. For example, an AI agent used for customer support that is powered by the GPT-4 LLM is able to interpret customer complaints, query backend systems using tool integration, and could then respond in a contextually appropriate and emotionally aware manner [RoTs25b].

Large Image Models (LIMs): such as CLIP [RaKi21] and BLIP-2 [LiLi23] DeepSeek Janus-Pro-7B extend the agent's capabilities into the visual domain. Trained on pairs of image-text, the LIMs can enable perception-based tasks including (a) image classification, (b) object detection, and (3) vision-language grounding. Such capabilities are becoming essential for agents operating in domains such as autonomous vehicles [ElAs25], [PaLe24], robotics [SoZh23], and visual content moderation [AhKh24].

Generative AI: In particular, the success of ChatGPT promoted the use of Generative Agents, which are LLM (Large Language Model) based systems, developed to produce outputs such as text,images, and program code from user prompts, i.e., explicit, visible input a user provides to an AI model to request a specific task or information) [LuAl24], [ZhCh24].

Such Generative AI agents were quite rapidly adopted across several applications, ranging from conversational assistants (e.g., GitHub Copilot [PeKa23]) and platforms for content generation (e.g., Jasper [LiLa19]) towards creative tools (e.g., Midjourney [Ja-Ro22]). These developments were considered as significant improvements in sectors, such as digital design, marketing, and software prototyping throughout 2023 and beyond.

The generative models are highly communicative, but they are having a reactive behavior, since they are able to produce output only when they are explicitly user prompted and are mainly not able to function autonomously or engage in self-initiated reasoning [GaBa23], [PePa24].

Some Key Characteristics of Generative AI are:

- **Reactivity**: Their operations are triggered by user-specified prompts and they as well, lack internal states, persistent memory, or goal/objective following mechanisms, see e.g., [LiWa23], [AlMi24].

- **Multi-modal Capability**: Modern generative systems can produce a diverse array of outputs, including coherent narratives, executable code, realistic images, and even speech transcripts

- **User Prompt Dependency and Statelessness**: Typical generative systems are stateless in that they do not retain context across interactions unless explicitly prompted, see e.g.,[DeLe24], [ChLe24]. Although, recent advancements like GPT-4.1 support larger context windows-up to 1 million tokens-and are better able to utilize that context enabled by the improved long-text comprehension, see e.g., [OpenAI25]. Furthermore, Generative AI design also lacks intrinsic feedback loops e.g., [PaJo24], state management, e.g., [Nabb24], or multi-step planning a requirement for autonomous decision-making and iterative goal refinement, e.g., [WeZh25].

Although Generative AI systems have a remarkable generative fidelity, such systems are constrained by their inability to act upon the environment or manipulate digital tools independently. Examples are: inability to (1) search the internet, (2) parse real-time data, nor (3) interact with APIs without human intervention or other supporting tools.

Currently, the AI landscape experienced a quite rapid transformation, evolving from the use of standalone LLMs toward more autonomous, and more task-oriented platforms [S'aCu24]. This evolution experienced two major post-generative phases: AI Agents and Agentic AI.

AI Agents: In particular, the limitations of generative AI on being constrained in acting upon the environment or manipulate digital tools independently, such as handling dynamic tasks, maintaining state continuity, or executing multi-step plans, have led to the development of AI tool-augmented systems, commonly referred to as AI Agents, see e.g. [BaAs23].

The AI agents build upon the language processing backbone of LLMs, but they support additional infrastructure such as memory buffers, tool-calling APIs, reasoning chains, and planning routines able to bridge the gaps between the passive response generation and successful active task completion. This architectural evolution marks a critical shift in AI system design: from content creation to autonomous task execution, see e.g., [LiDu25], [GuCh25]. The evolution from generative systems to AI Agents identifies a progressive layering of functionality, which ultimately enables the development of agentic behaviors.

Figure 53 shows the three key characteristics of AI Agents, which are:

- **Autonomy**: A key feature of AI Agents is their ability to function with minimal or no human intervention after their deployment [HaSc23].

- **Task-Specificity**: AI Agents are purpose-built for narrow, and well-defined tasks [Kris25], [PaSh25]. They are in particular, optimized to execute repeatable operations within a fixed domain, such as email filtering [EzSh24], [SiPa20], database querying [KhSa24], or calendar coordination [BuWi25], [Enda24].

- **Reactivity and Adaptation**: AI Agents often include basic mechanisms for interacting with dynamic inputs, allowing them to respond to real-time stimuli such as user requests, external API calls, or state changes in software environments [DeGu25], [RaMe24]. Some AI agent systems integrate basic learning capabilities [PaSk07] or integrate updated context buffers to refine behavior over time, particularly in settings like personalized recommendations or conversation flow management [KaSt24], [HuLi23], [BaAl22].
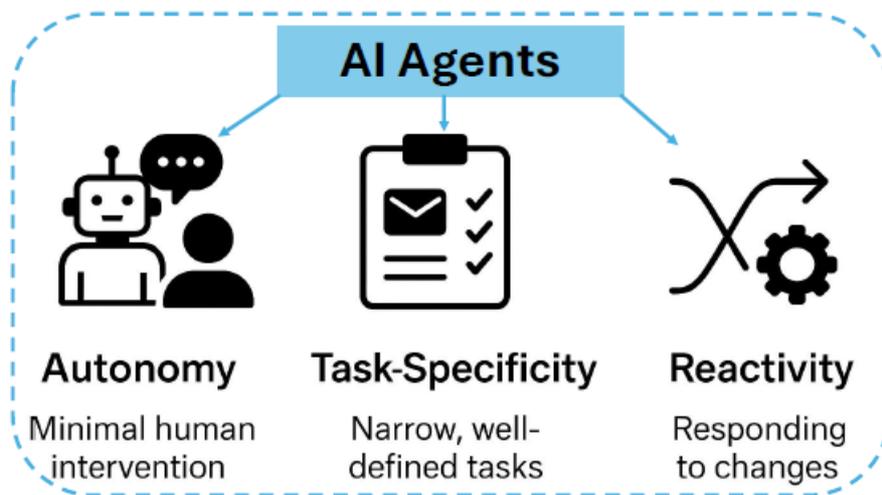
Figure 53: Key characteristics of AI Agents autonomy, task-specificity, and reactivity for agent design and operational behavior, copied from [SaRo25]

Agentic AI: As use cases increasingly require context retention, task interdependence, and as well adaptability across dynamic environments, the single AI agent model proves insufficient [PeLi24], [ShSh25]. In particular, although, AI Agents represent a significant improvement in artificial intelligence capabilities, such as in automating narrow tasks through tool-augmented reasoning, current research studies identify significant limitations that constrain their applicability. Such constraints are scalability in complex, dynamic, multi-step, and/or cooperative scenarios [WuYu23], [FeXu25]. The Agentic AI paradigm is a more advanced mechanism that can alleviate the AI Agent constraints. In particular, Agentic AI extends the capabilities of traditional AI Agents by enabling multiple specialized AI agents to collaborate and complete goals and tasks through collaborative reasoning and multi-step planning [NiLi25], structured communication see e.g., [ZhTa25], [MiRa25], shared memory [XuLi25], [RiCr25], and dynamic role assignment [AcKu25].

Currently, there are research and standardization activities (in e.g., 3GPP) on investigating how AI agents and Agentic AI systems can be applied and used in cellular (e.g., 6G) core networks, see e.g., [ChSu25], [LiSh25], [MoHo25], [FiAt22], [KhSa22], [LeGr22], [3GPP TR 22.870].

In particular, the AI Agent acts as a deterministic component with limited scope, while Agentic AI reflects distributed intelligence, characterized by goal decomposition, inter-agent communication, and contextual adaptation, demonstrating key characteristics of the modern agentic AI frameworks.

Examples of Agent to Agent communication protocols are:

- MCP (Model Context Protocol), see: https://github.com/modelcontextprotocol

- A2A (Agent to Agent) protocol, see: https://a2aprotocol.ai/ and https://a2aprotocol.ai/blog/impact-analysis-google-donating-a2a-protocol-linux-foundation

Key differences between AI agents and Agentic AI functionalities

An example of a conceptual illustration that shows the distinction between AI Agents and Agentic AI through the analogy of smart home systems, is shown in **Figure 54**.
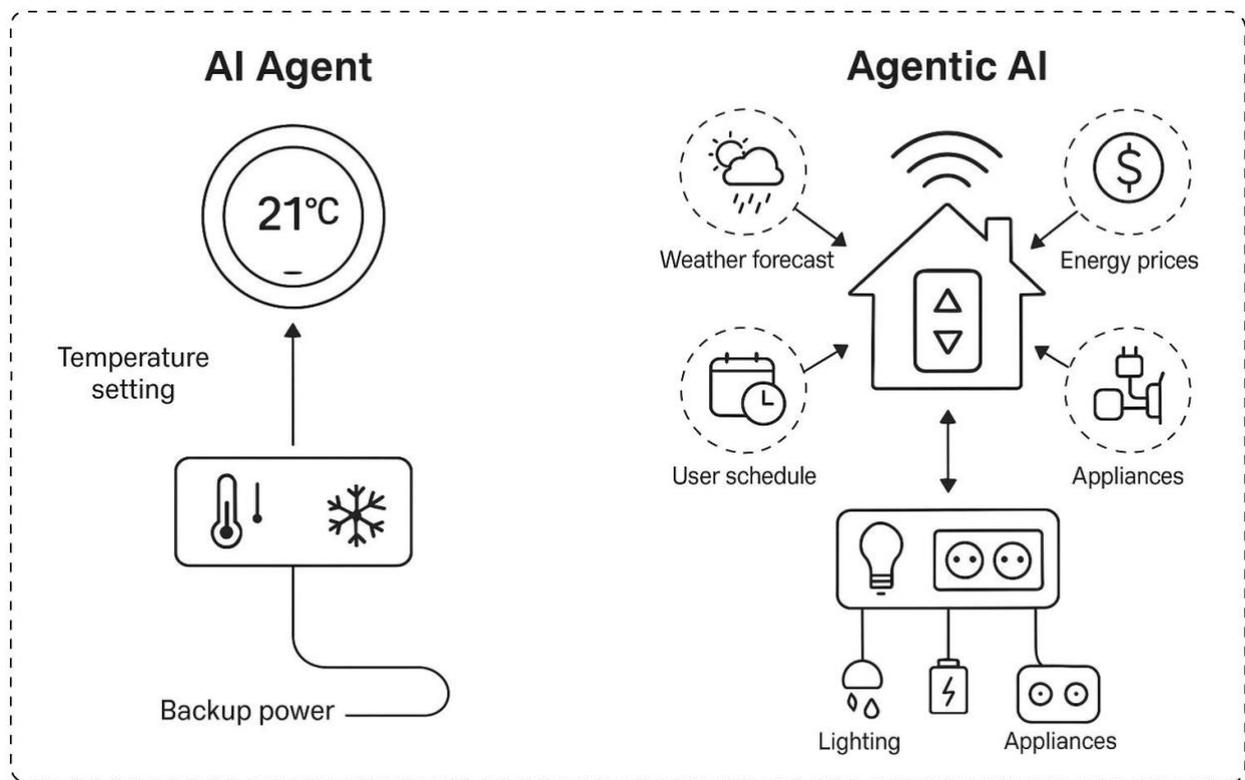
Figure 54: Comparative illustration of AI Agent vs. Agentic AI synthesizing conceptual distinctions. Left: A single-task AI Agent. Right A multi-agent Agentic AI system, copied from [SaRo25]

As depicted in **Figure 54**, the left part of the figure depicts a traditional AI Agent in the form of a smart thermostat.

This standalone AI agent receives a user-defined temperature setting and then autonomously controls the heating or cooling system to maintain the target temperature. While the right part of the same figure depicts an Agentic AI system that is embedded in a comprehensive smart home ecosystem. In this Agentic AI system, multiple specialized Ai agents interact synergistically in order to manage diverse aspects such as (1) user daily schedule, being aware about the user presence or absence (2) weather forecasting, (3) energy pricing optimization, (4) security monitoring, and (4) appliance management. Important to note that these agents are not just reactive modules. In particular, (1) they communicate dynamically, (2) share memory states, and (3) collaboratively align actions toward a high-level system goal (such as optimizing comfort, safety, and energy efficiency in real-time). As example, the AI agent that controls the weather forecast task might signal upcoming heatwaves, and in this way prompting the appliance management AI agent to start early pre-cooling via solar energy before peak pricing hours, using the information that was communicated previously, via the energy pricing AI agent). At the same time, the system might delay high-energy tasks (appliance management AI agent) or activate surveillance systems (during occupant absence, integrating decisions across domains.

Figure 54 illustrates that a single AI Agent acts as a deterministic component with limited scope, while Agentic AI reflects distributed intelligence, characterized by goal/task decomposition, inter-agent communication, and contextual adaptation, demonstrating key characteristics of the modern agentic AI frameworks.

While AI Agents and Agentic AI systems represent increasingly autonomous and interactive systems, both paradigms utilize generative architectures as their foundations, especially LLMs and LIMs and generative AI. The key differences between AI Agents and Agentic AI systems are provided in Table 25, in terms of scope, autonomy, architectural composition, coordination strategy, and operational complexity, which is based on the work published in [SaRo25] and derived from AutoGen [WuBa23] and ChatDev [QiLi23].

In particular, **Table 25** compares their definitions, levels of autonomy, capacity for handling task complexity, collaboration styles, learning and adaptation scope and typical application domains.

Table 25: Key Structural, Functional, and Operational Differences Between AI Agents and Agentic AI Systems, based on [SaRo25]

Feature	AI Agent	Agentic AI
Definition	Autonomous software programs that perform specific task	Systems of multiple AI agents collaborating to achieve complex goals
Autonomy Level	High autonomy within specific tasks.	Broad level of autonomy with the ability to manage multi-step, complex tasks and systems.
Task Complexity	Typically handle single, specific tasks.	Handle complex, multi-step tasks requiring coordination.
Collaboration	Operate independently.	Involve multi-agent information sharing, collaboration and cooperation.
Learning and Adaptation	Learn and adapt within their specific domain.	Learn and adapt across a wider range of tasks and environments.
Applications	Customer service chatbots, virtual assistants, automated workflows.	Supply chain management, business process optimization, virtual project managers.

3.9.1   Architectural evolution from traditional AI agents to Agentic AI systems

Agentic AI systems inherit the modularity of AI Agents but extend their architecture to support distributed intelligence, inter-agent communication, and iterative planning.

Figure 55, depicts the architectural evolution from traditional AI Agents to modern Agentic Ai systems. The left part of the figure shows the AI Agent key capabilities, such as Perception, Reasoning and Action, and expands, in the right part of the figure to the Agentic AI advanced capabilities, including Specialized Agents, Advanced Reasoning & Planning, Persistent Memory, and Orchestration. **Figure 55** depicts as well emergent properties such as Multi- Agent Collaboration, System Coordination, Shared Context, and Task Decomposition, and all are enclosed within a dotted boundary signifying layered modularity and the transition to distributed, adaptive agentic AI intelligence.
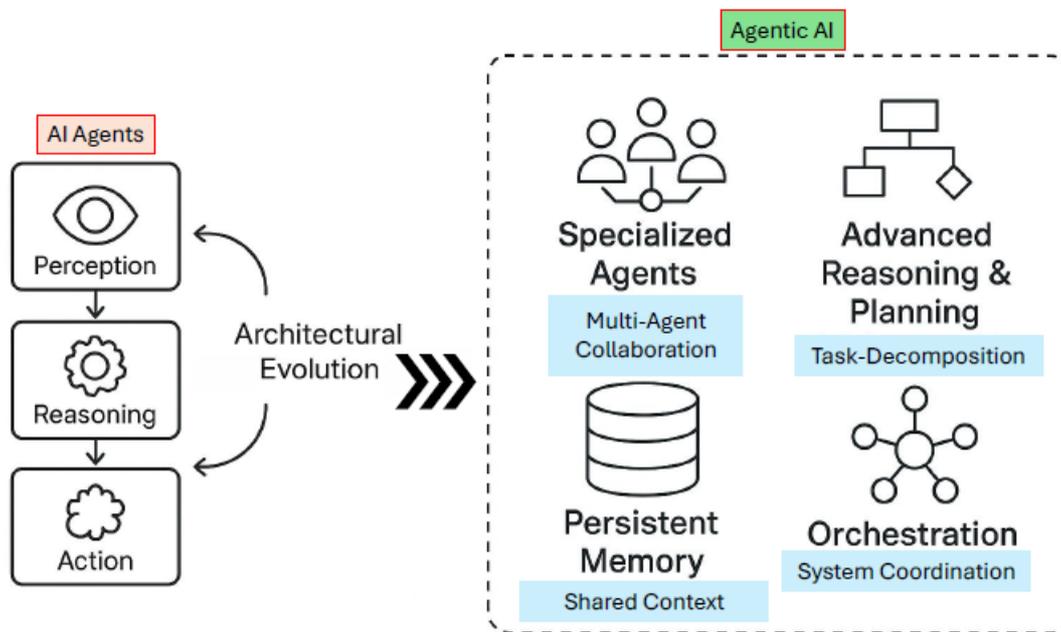
Figure 55: Architectural evolution from traditional AI Agents to modern Agentic Ai systems, copied from [SaRo25]

3.9.2 Examples of Standardisation activities on Agentic AI

Currently several standardization activities related Agentic AI are being investigated and some of them already started. The SDOs that are currently cover the Agentic AI topic are:

3GPP (examples):

- [3GPP 6G Work Items discussions on Agentic core network:](#) during a 3GPP TSG SA WG2 meeting that took place in Goteborg, Sweden in August 2025, a list 6G Work items focusing on Agentic core networks for 6G were discussed.

- 3GPP TR 22.870 "[Study on 6G Use Cases and Service Requirements](#)": a Technical Report that discusses among others Agentic AI 6G use cases.

ETSI (examples):

- ETSI GR ENI 051 V4.1.1 (2025-02), "[Study on AI Agents based Next-generation Network Slicing](#)"

IETF (examples):

- IETF draft-rosenberg-ai-protocols-00, "[Framework, Use Cases and Requirements for AI Agent](#)"

- IETF draft-stephan-ai-agent-6g-00, "[AI Agent protocols for 6G systems](#)"

- IETF draft-du-ai-agent-communication-6g-aspect-00, "[Use Cases and Requirements of AI Agent Communication from 6G Aspect](#)"

- IETF draft-yu-ai-agent-use-cases-in-6g-01, "[AI Agent Use Cases and Requirements in 6G Network](#)"

- IETF draft-akhavain-moussa-ai-network-00, "[AI Network for Training, Inference, and Agentic Interactions](#)"

- IETF draft-hw-ai-agent-6g-00, "[AI Network for Training, Inference, and Agentic Interactions](#)"

- IETF draft-campbell-agentic-http-00, "[A Best Current Practice for Agentic Interactions over HTTP](#)"

- IETF draft-wahl-scim-agent-schema-01, "System for Cross-domain Identity Management: Agentic Identity Schema"

- IETF draft-rosenberg-oauth-aauth-00, "AAuth - Agentic Authorization OAuth 2.1 Extension"

- IETF draft-huang-acme-scalable-agent-enrollment-00, "Extending Certificate Enrollment Protocols for Scalable Agentic AI Identity"

- IETF draft-huang-rats-agentic-eat-cap-attest-00, "Capability Attestation Extensions for the Entity Attestation Token (EAT) in Agentic AI Systems"

- IETF draft-zhao-nmop-network-management-agent-02, "AI based Network Management Agent(NMA): Concepts and Architecture"

- IETF draft-yue-anima-agent-recovery-networks-00, "Task-Oriented Multi-Agent Recovery Framework for High-Reliability in Converged Mobile Networks"

- IETF draft-liu-agent-context-protocol-00, "Agent Context Protocol"

- IETF draft-yang-ioa-protocol-00, "Internet of Agents Protocol (IoA Protocol) for Heterogeneous Agent Collaboration"

- IETF draft-jiang-cats-reference-acn-00, "CATS Reference Model for AI-Agent Communication Network"

- IETF draft-zhang-agent-gap-network-00, "Problem Statement and Gap Analysis for Agent-enabled Mobile Core Network

3.9.3 References

[SaRo25] R. Sapkota, K. I. Roumeliotis, M. Karkee "AI Agents vs. Agentic AI: A Conceptual Taxonomy, Applications and Challenges", Information Fusion, 2025, see: https://doi.org/10.1016/j.inffus.2025.103599

[RaHo25] T. Raheem; G. Hossain, "Agentic AI Systems: Opportunities, Challenges, and Trustworthiness", IEEE International Conference on Electro Information Technology (eIT) 2025, see: https://ieeexplore.ieee.org/abstract/document/11103638

[Pati25] A. Kumar Pati, "Agentic AI: A Comprehensive Survey of Technologies, Applications, and Societal Implications", IEEE Access, 2025, se: https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=11071266

[Sagr19]] Sagramsingh R. AI: a global survey. Washington, DC: IEEE-USA; 2019.

[AcKu25] D. B. Acharya, K. Kuppan, and B. Divya, "Agentic ai: Autonomous intelligence for complex goals–a comprehensive survey," IEEE Access, 2025.

[ChSu25] Ziqi Chen, Qi Sun, Nan Li, Xiang Li, Yang Wang, Chih-Lin, "Enabling Mobile AI Agent in 6G Era: Architecture and Key Technologies", IEEE Network, Vol. 38, Issue 5, pp 66-75, 2025, see: https://dl.acm.org/toc/ieeenetw/2024/38/5

[LiSh25] Xu Li, Weisen Shi, Hang Zhang, Chenghui Peng, Shaoyun Wu, Wen Tong, "The Agentic-AI Core: An AI-Empowered, Mission-Oriented Core Network for Next-Generation Mobile Telecommunications", Elsevier, Engineering 2025, see: https://www.sciencedirect.com/science/article/pii/S209580992500325X?via%3Dihub

[MoHo25] H.G. Moussa, A. Akhavain, S. M. Hosseini, B. McCormick, "Distributed Learning and Inference Systems: A Networking Perspective", IEEE Network, 2025, see: https://ieeexplore.ieee.org/document/11015802

[FiAt22] C. Fiandrino, G. Attanasio, M. Fiore, J. Widmer, "Toward native explainable and robust AI in 6G networks: Current state, challenges and road ahead," Computer Communications, Volume 193, 2022, see: https://www.sciencedirect.com/science/article/abs/pii/S0140366422002389

[KhSa22] L. U. Khan, W. Saad, D. Niyato, Z. Han and C. S. Hong, "Digital-Twin-Enabled 6G: Vision, Architectural Trends, and Future Directions," in IEEE Communications Magazine, vol. 60, no. 1, pp. 74-80, January 2022

[LeGr22] K.-D. Lee and C. Gray-Preston, "Everyday Living Assisted by 6G Applications and Solutions," IEEE Wireless Communications Magazine, October 2022. (Everyday Living Assisted by 6G Applications and Solutions | IEEE Journals & Magazine | IEEE Xplore

[3GPP TR 22.870] 3GPP TR 22.870 "Study on 6G Use Cases and Service Requirements", 3GPP Release 20, June 2025, see: https://www.3gpp.org/ftp/Specs/archive/22_series/22.870/22870-031.zip

[DeGu25] Z. Deng, Y. Guo, C. Han, W. Ma, J. Xiong, S. Wen, and Y. Xiang, "Ai agents under threat: A survey of key security challenges and future pathways," ACM Computing Surveys, vol. 57, no. 7, pp. 1–36, 2025.

[S'aCu24] J. S´anchez Cuadrado, S. P´erez-Soler, E. Guerra, and J. De Lara, "Automating the development of task-oriented llm-based chatbots,"Interfaces, pp. 1–10, 2024.

[LuAl24] Y. Lu, A. Aleta, C. Du, L. Shi, and Y. Moreno, "Llms and generative agent-based models for complex systems research," Physics of Life Reviews, 2024.

[ZhCh24] A. Zhang, Y. Chen, L. Sheng, X. Wang, and T.-S. Chua, "On generative agents in recommendation," in Proceedings of the 47th international ACM SIGIR conference on research and development in Information Retrieval, pp. 1807–1817, 2024.

[PeKa23] S. Peng, E. Kalliamvakou, P. Cihon, and M. Demirer, "The impact of ai on developer productivity: Evidence from github copilot," arXiv preprint arXiv:2302.06590, 2023.

[LiLa19] J. Li, V. Lavrukhin, B. Ginsburg, R. Leary, O. Kuchaiev, J. M. Cohen, H. Nguyen, and R. T. Gadde, "Jasper: An end-to-end convolutional neural acoustic model," arXiv preprint arXiv:1904.03288, 2019.

[Ja-Ro22] A. Jaruga-Rozdolska, "Artificial intelligence as part of future practices in the architect's work: Midjourney generative tool as part of a process of creating an architectural form," Architectus, no. 3 (71, pp. 95–104, 2022.

[YaYu23] H. Yang, S. Yue, and Y. He, "Auto-gpt for online decision making: Benchmarks and additional opinions," arXiv preprint arXiv:2306.02224, 2023, see: https://arxiv.org/abs/2306.02224

[BoWi25] P. Bornet, J. Wirtz, T. H. Davenport, D. De Cremer, B. Evergreen, P. Fersht, R. Gohel, S. Khiyara, P. Sund, and N. Mullakara, Agentic Artificial Intelligence: Harnessing AI Agents to Reinvent Business, Work and Life. Irreplaceable Publishing, 2025.

[HaSc23] A. I. Hauptman, B. G. Schelble, N. J. McNeese, and K. C. Madathil, "Adapt and overcome: Perceptions of adaptive autonomous agents for human-ai teaming," Computers in Human Behavior, vol. 138, p. 107451, 2023.

[Kris25] N. Krishnan, "Advancing multi-agent systems through model context protocol: Architecture, implementation, and applications," arXiv preprint arXiv:2504.21030, 2025.

[PaSh25] H. Padigela, C. Shah, and D. Juyal, "Ml-dev-bench: Comparative analysis of ai agents on ml development workflows," arXiv preprint arXiv:2502.00964, 2025.

[EzSh24] C. S. Eze and L. Shamir, "Analysis and prevention of ai-based phishing email attacks," Electronics, vol. 13, no. 10, p. 1839, 2024.

[SiPa20] D. Singh, V. Patel, D. Bose, and A. Sharma, "Enhancing email marketing efficacy through ai-driven personalization: Leveraging natural language processing and collaborative filtering algorithms," International Journal of AI Advancements, vol. 9, no. 4, 2020.

[KhSa24] R. Khan, S. Sarkar, S. K. Mahata, and E. Jose, "Security threats in agentic ai system," arXiv preprint arXiv:2410.14728, 2024.

[Enda24] C. G. Endacott, "Enacting machine agency when ai makes one's day: understanding how users relate to ai communication technologies for no. 4, p. zmae011, 2024.

[RaMe24] M. Raees, I. Meijerink, I. Lykourentzou, V.-J. Khan, and K. Papangelis, "From explainable to interactive ai: A literature review on current trends in human-ai interaction," International Journal of Human-Computer Studies, p. 103301, 2024.

[PaSk07] Z. Pawlak and A. Skowron, "Rudiments of rough sets," Information sciences, vol. 177, no. 1, pp. 3–27, 2007.

[KaSt24] S. Kapoor, B. Stroebl, Z. S. Siegel, N. Nadgir, and A. Narayanan, "Ai agents that matter," arXiv preprint arXiv:2407.01502, 2024.

[HuLi23] X. Huang, J. Lian, Y. Lei, J. Yao, D. Lian, and X. Xie, "Recommender ai agent: Integrating large language models for interactive recommendations," arXiv preprint arXiv:2308.16505, 2023.

[BaAl22] A. M. Baabdullah, A. A. Alalwan, R. S. Algharabat, B. Metri, and N. P. Rana, "Virtual agents and flow experience: An empirical examination of ai-powered chatbots," Technological Forecasting and Social Change, vol. 181, p. 121772, 2022.

[AcAd23] J. Achiam, S. Adler, S. Agarwal, L. Ahmad, I. Akkaya, F. L. Aleman, D. Almeida, J. Altenschmidt, S. Altman, S. Anadkat, et al., "Gpt-4 technical report," arXiv preprint arXiv:2303.08774, 2023.

[ChNa23] A. Chowdhery, S. Narang, J. Devlin, M. Bosma, G. Mishra, A. Roberts, P. Barham, H. W. Chung, C. Sutton, S. Gehrmann, et al., "Palm: Scaling language modeling with pathways," Journal of Machine Learning Research, vol. 24, no. 240, pp. 1–113, 2023.

[RoTs25a] K. I. Roumeliotis, N. D. Tselikas, and D. K. Nasiopoulos, "Llms for product classification in e-commerce: A zero-shot comparative study of gpt and claude models," Natural Language Processing Journal, vol. 11, p. 100142, 6 2025.

[RaKi21] A. Radford, J. W. Kim, C. Hallacy, A. Ramesh, G. Goh, S. Agarwal, G. Sastry, A. Askell, P. Mishkin, J. Clark, et al., "Learning transferable visual models from natural language supervision," in International conference on machine learning, pp. 8748–8763, PmLR, 2021.

[RoTs25b] K. I. Roumeliotis, N. D. Tselikas, and D. K. Nasiopoulos, "Think before you classify: The rise of reasoning large language models for consumer complaint detection and classification," Electronics 2025, Vol. 14, Page 1070, vol. 14, p. 1070, 3 2025

[LiLi23] J. Li, D. Li, S. Savarese, and S. Hoi, "Blip-2: Bootstrapping language image pre-training with frozen image encoders and large language models," in International conference on machine learning, pp. 19730– 19742, PMLR, 2023.

[SoZh23] S. Sontakke, J. Zhang, S. Arnold, K. Pertsch, E. Bıyık, D. Sadigh, C. Finn, and L. Itti, "Roboclip: One demonstration is enough to learn robot policies," Advances in Neural Information Processing Systems, vol. 36, pp. 55681–55693, 2023.

[ElAs25] M. Elhenawy, H. I. Ashqar, A. Rakotonirainy, T. I. Alhadidi, A. Jaber, and M. A. Tami, "Vision-language models for autonomous driving: Clip-based dynamic scene understanding," Electronics, vol. 14, no. 7, p. 1282, 2025.

[PaLe24] S. Park, M. Lee, J. Kang, H. Choi, Y. Park, J. Cho, A. Lee, and D. Kim, "Vlaad: Vision and language assistant for autonomous driving," in Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision, pp. 980 987, 2024.

[AhKh24] S. H. Ahmed, M. J. Khan, and G. Sukthankar, "Enhanced multimodal content moderation of children's videos using audiovisual fusion," arXiv preprint arXiv:2405.06128, 2024.

[WuBa23] Q. Wu, G. Bansal, J. Zhang, Y. Wu, B. Li, E. Zhu, L. Jiang, X. Zhang, S. Zhang, J. Liu, et al., "Autogen: Enabling next-gen llm applications via multi-agent conversation," arXiv preprint arXiv:2308.08155, 2023, see: https://www.researchgate.net/publication/373164024_AutoGen_Enabling_Next-Gen_LLM_Applications_via_Multi-Agent_Conversation_Framework

[GaBa23] L. Gabora and J. Bach, "A path to generative artificial selves," in EPIA Conference on Artificial Intelligence, pp. 15–29, Springer, 2023.

[PePa24] G. Pezzulo, T. Parr, P. Cisek, A. Clark, and K. Friston, "Generating meaning: active inference and the scope and limits of passive ai," Trends in Cognitive Sciences, vol. 28, no. 2, pp. 97–112, 2024.

[LiWa23] X. Liu, J. Wang, J. Sun, X. Yuan, G. Dong, P. Di, W. Wang, and D. Wang, "Prompting frameworks for large language models: A survey," arXiv preprint arXiv:2311.12785, 2023.

[AlMi24] K. Alizadeh, S. I. Mirzadeh, D. Belenko, S. Khatamifard, M. Cho, C. C. Del Mundo, M. Rastegari, and M. Farajtabar, "Llm in a flash: Efficient large language model inference with limited memory," in Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), pp. 12562–12584, 2024.

[DeLe24] P. Denny, J. Leinonen, J. Prather, A. Luxton-Reilly, T. Amarouche, B. A. Becker, and B. N. Reeves, "Prompt problems: A new programming exercise for the generative ai era," in Proceedings of the 55th ACM Technical Symposium on Computer Science Education V. 1, pp. 296–302, 2024.

[ChLe24] C. Chen, S. Lee, E. Jang, and S. S. Sundar, "Is your prompt detailed enough? exploring the effects of prompt coaching on users' perceptions, engagement, and trust in text-to-image generative ai tools," in Proceedings of the Second International Symposium on Trustworthy Autonomous Systems, pp. 1–12, 2024.

[OpenAI25] OpenAI, "Introducing gpt-4.1 in the api," 4 2025.

[PaJo24] A. Pan, E. Jones, M. Jagadeesan, and J. Steinhardt, "Feedback loops with language models drive in-context reward hacking," arXiv preprint arXiv:2402.06627, 2024.

[Nabb24] K. Nabben, "Ai as a constituted system: accountability lessons from an llm experiment," Data & policy, vol. 6, p. e57, 2024.

[WeZh25] H. Wei, Z. Zhang, S. He, T. Xia, S. Pan, and F. Liu, "Plangenllms: A modern survey of llm planning capabilities," arXiv preprint arXiv:2502.11221, 2025.

[BaAs23] A. Bandi, P. V. S. R. Adapa, and Y. E. V. P. K. Kuchi, "The power of generative ai: A review of requirements, models, input–output formats, evaluation metrics, and challenges," Future Internet, vol. 15, no. 8, p. 260, 2023.

[LiDu25] Y. Liu, H. Du, D. Niyato, J. Kang, Z. Xiong, Y. Wen, and D. I. Kim, "Generative ai in data center networking: Fundamentals, perspectives, and case study," IEEE Network, 2025.

[GuCh25] C. Guo, F. Cheng, Z. Du, J. Kiessling, J. Ku, S. Li, Z. Li, M. Ma, T. Molom-Ochir, B. Morris, et al., "A survey: Collaborative hardware and software design in the era of large language models," IEEE Circuits and Systems Magazine, vol. 25, no. 1, pp. 35–57, 2025.

[NiLi25] L. Ning, Z. Liang, Z. Jiang, H. Qu, Y. Ding, W. Fan, X.-y. Wei, S. Lin, H. Liu, P. S. Yu, et al., "A survey of webagents: Towards next-generation ai agents for web automation with large foundation models," arXiv preprint arXiv:2503.23350, 2025.

[WuYu23] Z. Wu, C. Yu, C. Chen, J. Hao, and H. H. Zhuo, "Models as agents: Optimizing multi-step predictions of interactive local models in modelbased multi-agent reinforcement learning," in Proceedings of the AAAI Conference on Artificial Intelligence, vol. 37, pp. 10435–10443, 2023.

[FeXu25] Z. Feng, R. Xue, L. Yuan, Y. Yu, N. Ding, M. Liu, B. Gao, J. Sun, and G. Wang, "Multi-agent embodied ai: Advances and future directions," arXiv preprint arXiv:2505.05108, 2025.

[ZhTa25] R. Zhang, S. Tang, Y. Liu, D. Niyato, Z. Xiong, S. Sun, S. Mao, and Z. Han, "Toward agentic ai: generative information retrieval inspired intelligent communications and networking," arXiv preprint arXiv:2502.16866, 2025.

[MiRa25] E. Miehling, K. N. Ramamurthy, K. R. Varshney, M. Riemer, D. Bouneffouf, J. T. Richards, A. Dhurandhar, E. M. Daly, M. Hind, P. Sattigeri, et al., "Agentic ai needs a systems theory," arXiv preprint arXiv:2503.00237, 2025.

[XuLi25] W. Xu, Z. Liang, K. Mei, H. Gao, J. Tan, and Y. Zhang, "A-mem: Agentic memory for llm agents," arXiv preprint arXiv:2502.12110, 2025.

[RiCr25] C. Riedl and D. De Cremer, "Ai for collective intelligence," Collective Intelligence, vol. 4, no. 2, p. 26339137251328909, 2025.

[PeLi24] L. Peng, D. Li, Z. Zhang, T. Zhang, A. Huang, S. Yang, and Y. Hu, "Human-ai collaboration: Unraveling the effects of user proficiency and ai agent capability in intelligent decision support systems," International Journal of Industrial Ergonomics, vol. 103, p. 103629, 2024.

[ShSh25] H. Shirado, K. Shimizu, N. A. Christakis, and S. Kasahara, "Realism drives interpersonal reciprocity but yields to ai-assisted egocentrism in a coordination experiment," in Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems, pp. 1–21, 2025.

[QiLi23] C. Qian, W. Liu, H. Liu, N. Chen, Y. Dang, J. Li, C. Yang, W. Chen, Y. Su, X. Cong, et al., "Chatdev: Communicative agents for software development," arXiv preprint arXiv:2307.07924, 2023, see: https://www.researchgate.net/publication/372416377_Communicative_Agents_for_Software_Development

[HoZh23] S. Hong, X. Zheng, J. Chen, Y. Cheng, J. Wang, C. Zhang, Z. Wang, S. K. S. Yau, Z. Lin, L. Zhou, et al., "Metagpt: Meta programming for multi-agent collaborative framework," arXiv preprint arXiv:2308.00352, vol. 3, no. 4, p. 6, 2023, see: https://arxiv.org/abs/2308.00352

[WaXi23] G. Wang, Y. Xie, Y. Jiang, A. Mandlekar, C. Xiao, Y. Zhu, L. Fan, and A. Anandkumar, "Voyager: An open-ended embodied agent with large language models," arXiv preprint arXiv:2305.16291, 2023., see: https://arxiv.org/abs/2305.16291

[LiHa23] G. Li, H. Hammoud, H. Itani, D. Khizbullin, and B. Ghanem, "CAMEL: Communicative Agents for "Mind" Exploration of Large Language Model Society" pp. 51991–52008, arXiv:2303.17760, 2023 https://arxiv.org/abs/2303.17760

[ReZo22] S. Reed, K. Zolna, E. Parisotto, S. G. Colmenarejo, A. Novikov, G. Barth-Maron, M. Gimenez, Y. Sulsky, J. Kay, J. T. Springenberg, et al., "A generalist agent," arXiv preprint arXiv:2205.06175, 2022, see: https://arxiv.org/abs/2205.06175

4. Conclusions and Recommendations

It is expected that 5G and beyond 5G systems will extend mobile communication services beyond mobile telephony, mobile broadband, and massive machine-type communication into new application domains, so-called vertical domains.

[AIOTI-IoT-relation-5G] highlighted specific IoT vertical domain use cases and determined the specific requirements they impose on the network infrastructure. This report highlights additional IoT and Edge Computing vertical domain use cases collected by AIOTI and determines the specific requirements they impose on the underlying 5G and Beyond 5G network infrastructure. These use cases and requirements can be used by Standards Developing Organizations (SDOs), such as 3GPP, ITU-T, ISO and IEEE as requirements for automation in vertical domains focusing on critical communications. In addition to these use cases also emerging topics in the area of 5G/6G technology are as well introduced.

In particular, this report lists first relevant IoT and edge computing use cases and their possible requirements on an underlying 5G/6G communication infrastructure.

The Release 2.0 of this report included 6 additional use cases in the areas of: (1) use of drones, (2) 5G cloud-RAN, (3) Health-Critical Remote Operations, (4) preliminary 6G use cases. Secondly, emerging topics in the context of the Beyond 5G communication infrastructure, relevant for IoT and edge computing use cases are identified.

The Release 3.0 of this report includes 6 additional use cases as well in the area of Edge-Cloud Orchestration in Section 2.13.

The Release 4.0 of this report includes 14 additional use cases in the areas of: (1) Digital Twin, (2) autonomous urban transportation, (3) critical Infrastructure support applications (smart health and connected vehicles), (4) preliminary 6G use cases, (5) use of drones, (6) smart manufacturing and automation, (7) service trust and liability management, (8) Edge-Cloud orchestration and (9) smart agriculture.

The Release 5.0 of this report, compared to Release 4.0, introduced the following changes:

7. the use cases, developed in projects that are: (1) completed before 2025 and (2) were included in Release 4.0 of the report, are not anymore included in this release (Release 5.0) of the report, but they can be retrieved via Section 2 of AI, IoT and Edge Continuum impact and relation on 5G/6G: enabling technologies and challenges Release 4.0.

- in Section 3, added a subsection on Key AI concepts

- in Section 2, a list of additional 12 use cases are introduced in the areas of (1) Smart Manufacturing and Automation, (2) Preliminary 6G use cases and (3) Edge-Cloud orchestration.

4.1 Requirements

By analysing the requirements that are derived from the presented use cases as described in the Section 2, it can be concluded that for these use cases the requirements listed in [Networld2020-SRIA] report, see Annex III as well are covering the needs that each of these use cases impose on the underlying 5G/6G infrastructure.

In particular, requirements are identified by these use cases added in this Release 5 of the report can be found via the sections referenced below.

2.1 Smart Manufacturing and Automation

- 2.1.1   Factory-in-a-Box: Multiprotocol Beyond-5G Testbed for Smart Manufacturing

Requirements derived in this use case can be found in: _**Section 2.1.1.10 Potential Requirements**_.

2.2 Preliminary 6G Use Cases

- 2.2.1 6G-BRICKS (6G- Building Reusable testbed Infrastructures for validating Cloud-to-device breakthrough technologies) Use Cases
    - 2.2.1.1 Use Case 1: Metaverse as an enabler of a Modern Workplace;
        - Requirements derived in this use case can be found in:
            - _Section 2.2.1.1.10 Potential Requirements for PoC1: Holo-conferencing in a Virtual Meeting room_
            - _Section 2.2.1.1.11 Potential Requirements for PoC 2 Multiuser Team Building Activities with Advanced Full Volumetric Capture and In-cloud Media Processing_
    - 2.2.1.2 Use Case 2: 6G applications for Industry 4.0
        - Requirements derived in this use case can be found in:
            - _Section 2.2.1.2.1.1 Potential Requirements for PoC 1: Autonomous robots in Industry 4.0 (viewers/participants quality of experience_)
            - _Section 2.2.1.2.2.1 Potential Requirements for PoC 2: AR inspection of Industry 4.0 digital twin on site (for viewers/participants quality of experience)_
- 2.2.2 6G-DALI (6G DAta and ML operations automation via an end-to-end AI framework) Use Cases
    - 2.2.2.1 Use Case 1: Data Management and Experiment-on-Demand
        - Requirements derived in this use case can be found in:
            - _Section 2.2.2.1.10 Potential Requirements_
    - 2.2.2.2 Use Case 2: AIaaS for CDN Apps via Cross-Testbed Decentralized MLOps
        - Requirements derived in this use case can be found in:
            - _Section 2.2.2.2.10 Potential Requirements_
    - 2.2.2.3 Use Case 3: DTT and RLOps for Large and Medium-Scale Experiments
        - Requirements derived in this use case can be found in:
            - _Section 2.2.2.3.10 Potential Requirements_
- 2.2.3 6G-INTENSE (Intent-driven NaTive AI architecturE supporting Compute-Network abstraction and Sensing at the Deep Edge) Use Cases
    - 2.2.3.1 Use Case 1: Distributed Continuum towards Pervasive Computing
        - Requirements derived in this use case can be found in:
            - _Section 2.2.3.1.1.10 Potential Requirements for Use Case 1 - Scenario 1: Pervasive Computing in a distributed continuum_
            - _Section 2.2.3.1.2.10 Potential Requirements for Use Case 1 - Scenario 2: Edge Intelligence and Compute Interconnection_
    - 2.2.3.2 Use Case 2: Metaverse
        - Requirements derived in this use case can be found in:
            - _Section 2.2.3.2.1.10 Potential Requirements for Use Case 2 - Scenario 1: Joint Communication and Sensing for Optimal User Tracking in the Metaverse_

- *Section 2.2.3.2.2.10 Potential Requirements for Use Case 2 - Scenario 2: Fully Autonomous Metaverse FCAPS, sensing and Continuum Abstraction*
- 2.2.4 SUNRISE-6G (SUstainable federatioN of Research Infrastructures for Scaling-up Experimentation in 6G) Use Cases
  - o 2.2.4.1 Use Case 1 : Federated Metaverse
    - Requirements derived in this use case can be found in:
      - *Section 2.2.4.1.10 Potential Requirements*
  - o 2.2.4.2 Use Case 2: Collaborative Robotic Exploration of Mining Environments
    - Requirements derived in this use case can be found in:
      - *Section 2.2.4.2.10 Potential Requirements*
  - o 2.2.4.3 Use Case 3: Mobile Industrial Assets aided by NTN Edge Cloud & Communications
    - Requirements derived in this use case can be found in:
      - *Section 2.2.4.3.10 Potential Requirements*

2.4 Edge-Cloud Orchestration

- 2.4.1   COGNIT - CCAM services enhanced with AI assisted CEI framework
  - o Requirements derived in this use case can be found in:
    - Section *2.4.2.10 Potential Requirements*

4.2.   Emerging Topics

The following emerging topics that are related to IoT and Edge Computing that can impact the specifications and deployments of beyond 5G communication infrastructure, are identified:

1. Digital Twin (DT)
2. Deep Edge, Terminal and IoT Device Integration in B5G communication infrastructure
3. Edge, Mobile Edge Computing and Processing
4. Network and Server security for edge and IoT
5. Plug and Play Integrated Satellite and Terrestrial Networks
6. Autonomous and Hyper-connected On-demand Urban Transportation
7. Opportunities for IoT Components and Devices
8. EU legislative framework
9. Key AI concepts

For each of these emerging topics an overview and as well challenges are identified and briefly explained in Chapter 3.

ANNEX I   References

AIOTI-IoT-relation-5G] "IoT Relation and Impact on 5G", AIOTI, Release 3.0, April 2020, to retrieved via (accessed on 23 July 2021): https://aioti.eu/wp-content/uploads/2020/05/AIOTI-IoT-relation-and-impact-on-5G-R3-Published.pdf

[3GPP-TSG-RAN89E] 3GPP TSG RAN#89E, RP-201702: https://www.3gpp.org/ftp/TSG_RAN/TSG_RAN/TSGR_89e/Docs/RP-201702.zip

[5GPPP-Vision] 5G Vision, The 5G Infrastructure Public Private Partnership: the next generation of communication networks and services, 5GPPP, February 2015, to be retrieved via: https://5g-ppp.eu/wp-content/uploads/2015/02/5G-Vision-Brochure-v1.pdf

[3GPP TR 22.804] 3GPP TR 22.804, "Study on Communication for Automation in Vertical domains", Online: http://www.3gpp.org/DynaReport/22804.htm, 2018.

[5GPPP-verticals] 5G-PPP, "5G Empowering Vertical Industries," 02 2016. [Online]. Available: https://5g-ppp.eu/wp-content/uploads/2016/02/BROCHURE_5PPP_BAT2_PL.pdf.

[b-3GPP TR 26.918] Technical Report 3GPP TR 26.918 V16.0.0 (2018), 3rd Generation Partnership Project; Technical specification group services and system aspects; Virtual reality (VR) media services over 3GPP (Release 16).

[b-ETSI TR 126 928] "Extended Reality (XR) in 5G", 3GPP TR 26.928 version 16.1.0 Release 16, Jan 2021, to be retrieved via: https://www.etsi.org/deliver/etsi_tr/126900_126999/126928/16.01.00_60/tr_126928v160100p.pdf

[CiNe19] C. Cimino, E. Negri, L. Fumagalli, "Review of Digital Twin applications in manufacturing", Computers in Industry, 2019, 113, p.103130.

[Glaes12] E. S. D. Glaessgen, "The Digital Twin Paradigm for Future NASA and U.S. Air Force Vehicles," in 53rd AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Materials Conference - Special Session on the Digital Twin, Honolulu, HI, 2012.

[GaRo12] M. Garetti, P. Rosa, S. Terzi, „Life Cycle Simulation for the design of Product–Service Systems," Computers in Industry,, Elsevier, pp. 361-369, 2012.

[ErLi17] Ericsson and Arthur D. Little, "The 5G business potential", second edition, October 2017.

[ESA ESSB HB –U 002] ESA Space Debris Mitigation Compliance Verification Guidelines ESSB-HB-U-002: https://copernicus-masters.com/wp-content/uploads/2017/03/ESSB-HB-U-002-Issue119February20151.pdf

[Ericsson20] Ericsson Mobility Report: https://www.ericsson.com/4adc87/assets/local/mobility-report/documents/2020/november-2020-ericsson-mobility-report.pdf [ISO 2413] ISO 24113:2019 Space Systems-Space Debris Mitigation Requirements: https://www.iso.org/standard/72383.html

[Evans11] D. Evans, "The Internet of Things: How the Next Evolution of the Internet is Changing Everything," [Online]. Available: https://www.cisco.com/c/dam/en_us/about/ac79/docs/innov/IoT_IBSG_0411FINAL.pdf

[ISTINCT] Reference: Assessing satellite-terrestrial integration opportunities in the 5G environment: European Space Agency ARTES 1 Project "INSTINCT: Scenarios for Integration of Satellite Components in Future Networks' Contract No.: 4000110994/14/NL/AD

[ITU-T Y.3106] Recommendation ITU-T Y.3106 (2019), Quality of service functional requirements for the IMT-2020 network.

[ITU-T Y.3107] Recommendation ITU-T Y.3107 (2019), Functional architecture for QoS assurance management in the IMT-2020 network.

[ITU-T G.1035] Recommendation ITU-T G.1035 (2020), Influencing factors on quality of experience for virtual reality services.

[ITU-T Y.3102] Recommendation ITU-T Y.3102 (2018), Framework of the IMT-2020 network.

[ITU-T Y.3104] Recommendation ITU-T Y.3104 (2018), Architecture of the IMT-2020 network.

[ITU-T SG13 Y.3109] ITU-T SG13 Y.3109 (formerly Y.qos-ec-vr-req) "Quality of service assurance-related requirements and framework for virtual reality delivery using mobile edge computing supported by IMT-2020", published in April 2021 (https://www.itu.int/rec/T-REC-Y.3109-202104-I).

[ITU-T H.264] Recommendation ITU-T H.264 (2019), Advanced video coding for generic audiovisual services.

[ITU-T H.265] Recommendation ITU-T H.265 (2019), High efficiency video coding.

[ITU-T H.266] Recommendation ITU-T H.266 (2020), Versatile video coding.

[ITU-T E.860] Recommendation ITU-T E.860 (2002), Framework of a service level agreement.

[ITU-T SG13 Y.3109] ITU-T SG13 Y.3109 (formerly Y.qos-ec-vr-req) "Quality of service assurance-related requirements and framework for virtual reality delivery using mobile edge computing supported by IMT-2020", published in April 2021 (https://www.itu.int/rec/T-REC-Y.3109-202104-I).

[ISO/IEC TR 22417:2017] "Information technology — Internet of things (IoT) use cases", ISO/IEC TR 22417, November, 2017, see: https://www.iso.org/standard/73148.html

[ITU-R M.2410-0] International Telecommunications Union Radiocommunication Sector (ITU-R), "Minimum requirements related to technical performance for IMT-2020 radio interface(s)", Report ITU-R M.2410-0 (11/2017), November 2017, Online: https://www.itu.int/pub/R-REP-M.2410-2017.

[JML20] J. &. A. M. &. M. M. Lee, "5G and Smart Manufacturing," 2020.

[KaWa13] H. Kagermann, W. Wahlster and J. Helbig (Eds.), "Recommendations for implementing the strategic initiative Industrie 4.0: Final report of the Industrie 4.0 Working Group", 2013.

[KrKa18] W. Kritzinger, M. Karner, G. Traar, J.Henjes, W. Sihn „Digital Twin in manufacturing: A categorical literature review and classification," IFAC-PapersOnLine., 51 (2018), pp. 1016-1022, 2018.

[KoLa20] Kodheli, O., Lagunas, E., Maturo, N., Sharma, S.K., Shankar, B., Montoya, J.F.M., Duncan, J.C.M., Spano, D., Chatzinotas, S., Kisseleff, S. and Querol, J., 2020. Satellite communications in the new space era: A survey and future challenges. IEEE Communications Surveys & Tutorials.

[LeBa15] J. Lee, B. Bagheri, and H.-A. Kao, "A cyber-physical systems architecture for industry 4.0-based manufacturing systems, Manufacturing Letters", vol. 3, 2015, pp. 18-23.

[LeAz20] Jay Lee, Moslem Azamfar, Jaskaran Singh, Shahin Siahpour, "Integration of digital twin and deep learning in cyber-physical systems: towards smart manufacturing", IET Collab. Intell. Manuf., 2020, Vol. 2 Iss. 1, pp. 34-36

[LiGe19] Liolis, K., Geurtz, A., Sperber, R., Schulz, D., Watts, S., Poziopoulou, G., Evans, B., Wang, N., Vidal, O., Tiomela Jou, B. and Fitch, M., 2019. Use cases and scenarios of 5G integrated satellite-terrestrial networks for enhanced mobile broadband: The SaT5G approach. International Journal of Satellite Communications and Networking, 37(2), pp.91-112.

[MuBo22] M. Mukhiddinov, A. Bobomirzaevich Abdusalomov, J. Cho, "A Wildfire Smoke Detection System Using Unmanned Aerial Vehicle Images Based on the Optimized YOLOv5", Special Issue Advanced Computational Intelligence for Object Detection, Feature Extraction and Recognition in Smart Sensor Environments 2022-2023), 1 December 2022

[NASA-cubesats] https://www.nasa.gov/mission_pages/cubesats/overview

[Networld2020-SRIA] "Smart Networks in the context of NGI", SNS SRIA, Networld2020, September 2020, https://bscw.5g-ppp.eu/pub/bscw.cgi/d367342/Networld2020%20SRIA%202020%20Final%20Version%202.2%20.pdf

[Siemens2016] Siemens AG, "5G communication networks: Vertical industry requirements," 11 2016, to be retrieved via (accessed on 23 July 2021): http://www.virtuwind.eu/_docs/Siemens_PositionPaper_5G_2016.pdf.

[Satell-market] http://satellitemarkets.com/satellite-iot-game-changer-industry

[SiBa23] Radheshyam Singh, Kalpit Dilip Ballal, Michael Stübert Berger, Lars Dittmann, "Overview of Drone Communication Requirements in 5G", Lecture Notes in Computer Science book series (LNCS,volume 13533), 01 January 2023.

[TaQi19] F. Tao, Q. Qi, L. Wang, AYC. Nee, „Digital Twins and Cyber–Physical Systems toward Smart Manufacturing and Industry 4.0: Correlation and Comparison," Engineering. 5. , pp. 653-661, 2019.

[TaCa19] G. Tavola, A. Caielli and M. Taisch, "An "Additive" Architecture for Industry 4.0 Transition of Existing Production Systems," in STUDIES IN COMPUTATIONAL INTELLIGENCE, Springer, 2019, pp. 258-269.

ANNEX II   Template used for Use Case description

X. Use Case (title)

X.1 Description

- Provide motivation of having this use case, e.g., is it currently applied and successful; what are the business drivers, e.g., several stakeholder types will participate and profit from this use case

- Provide on a high level, the operation of the use case, i.e., which sequence of steps are used in this operation?

X.2 Source

- Provide reference to project, SDO, alliance, etc.

X.3 Roles and Actors

- Roles: Roles relating to/appearing in the use case

    o Roles and responsibilities in this use case, e.g., end user, vertical industry, Communication Network supplier/provider/operator, IoT device manufacturer, IoT platform provider, Insurance company, etc.

    o Relationships between roles

- Actors: Which are the actors with respect to played roles

- A detailed definition of the Roles and Actors is provided in [7].

X.4 Pre-conditions

- What are the pre-conditions that must be valid (be in place) before the use case can become operational

X.5 Triggers

- What are the triggers used by this use case

X.6 Normal Flow

- What is the normal flow of exchanged data between the key entities used in this use case: devices, IoT platform, infrastructure, pedestrians, vehicles, etc?

X.7 Alternative Flow

- Is there an alternative flow

X.8 Post-conditions

- What happens after the use case is completed

X.9 High Level Illustration

- High level figure/picture that shows the main entities used in the use case and if possible, their interaction on a high level of abstraction

X.10 Potential Requirements

This section should provide the potential requirements and in particular the requirements imposed towards the underlying communication technology

These requirements can be split in:

- Functional requirements

(to possibly consider them – but not limited to – with respect to the identified functions/capabilities)

- Non-functional requirements – possible consideration includes:

    o Flexibility

    o Scalability

    o Interoperability

    o Reliability

    o Safety

    o Security and privacy

    o Trust

As example of the format of such requirements is provided in Annex III and Annex IV.

X.11 Radio Specific requirements

X.11.1 Radio Coverage

- Radio cell range

Specification of expected maximum and typical radio ranges (indicate if LOS/NoLOS)

- Does the radio link crosses public spaces? Or is it constrained to indoor or customer premises?

- Is Multicell required?

(If YES, specify the required scope of the multicell arrangement. I.e. "building", "city", "global")

- Is handover required? Seamless? Tolerable impact in delay and jitter?

- Mobility: maximum relative speed of UE/FP peers

- Special coverage needs: i.e., maritime, aerial

X.11.2 Bandwidth requirements

- Peak data rate

- Average data rate

- Is traffic packet mode or circuit mode?

    o If circuit mode, is isochronicity required?

X.11.3 URLLC requirements

- Required Latency

(specify if it is one way or roundtrip)

- Required Reliability

(i.e., 99,99999%)

- Maximum tolerable jitter

X.11.4 Radio regimens requirements

- Desired and acceptable radio regimens (describe the desired and acceptable radio regimens

(i.e.: licensed - public mobile, licensed – specific license, license-exempt)

X.11.5 Other requirements

- UE power consumption

    o Rechargeable or primary battery?

    o Acceptable battery life

- Is terminal location required? location accuracy?

ANNEX III   KPIs defined in Networld2020[64] (SNS) Strategic and Research Innovation Agenda (SRIA) 2021 – 2027

Selected KPIs Forecast for Terrestrial Radio Communications during the short, medium, and long -term evolution of 5G NR.

Target KPI	5G NR (Rel.16) 2020	Short-term Evo ~2025	Medium-term Evo ~2028	Long-term Evo ~2030
Spectrum	<52.6 GHz	<150 GHz	<300 GHz	<500 GHz
Bandwidth	<0.5 GHz	<2.5 GHz	<5 GHz	<10 GHz
Peak Data Rate	DL: >20 Gbps UL: >10 Gbps	DL: >100 Gbps UL: >50 Gbps	DL: >200 Gbps UL: >100 Gbps	DL: >400 Gbps UL: >200 Gbps
User Data Rate	DL: >100 Mbps UL: >50 Mbps	DL: >500 Mbps UL: >250 Mbps	DL: >1 Gbps UL: >0.5 Gbps	DL:>2 Gbps UL: >1 Gbps
Density	>1 device/sqm	>1.5 device/sqm	>2 device/sqm	>5 device/sqm
Reliability [BLER]	URLLC: $>1\text{-}10^{-5}$	$>1\text{-}10^{-6}$	$>1\text{-}10^{-7}$	$>1\text{-}10^{-8}$
U-Plane Latency	URLLC: <1 ms	<0.5 ms	<0.2 ms	<0.1 ms
C-Plane Latency	<20 ms	<10 ms	<4 ms	<2 ms
Energy Efficiency (Network/Terminal)	Qualitative	>30 % gain vs IMT-2020	>70 % gain vs IMT-2020	>100% gain vs IMT-2020
Mobility	<500 Km/h	<500 Km/h	<500 Km/h	<1000 Km/h
Positioning accuracy	NA (<1 m)	<30 cm	<10 cm	<1 cm

[Table copied from [Networld2020-SRIA] - Networld 2020 (SNS) Strategic and Research Innovation Agenda (SRIA) 2021 – 2027, see: https://bscw.5g-ppp.eu/pub/bscw.cgi/d367342/Networld2020%20SRIA%202020%20Final%20Version%202.2%20.pdf]

---

[64] Networld2020 ETP has been renamed to NetwolrdEurope ETP, see: https://www.networldeurope.eu

Selected KPIs Forecast for Satellite Radio Communications during the short, medium, and long-term evolution of 5G NR

KPI	Short Term Evo	Medium-Term Evo	Long-Term Evo
Minimization of unmet capacity[1]	<0.1.%	<0.05%	<0.01%
Maximization of satellite resource utilization[2]	>99%	>99.9%	>99.99%
Time to reallocate satellite resources[3]	<1 min	<5 sec	<1 sec
Solving and detecting time of satellite operation incidents	<10 min	<5min	< 1 min
Energy Reduction using adaptive intersegment links	>50%	>80%	>90%
Connectivity gain for converged satellite cloud scenarios[4]	>100%	>150%	>200%
Reduction of required manual intervention[5]	>50%	>80%	>90%
Widespread IoT coverage [6]	> 50%	>99%	> 99.9%
Reliability (perceived zero downtime)[7]	>50%	>99%	>99.9%
Experienced data rate (Broadband)	DL: >50 Mbit/s UL: >25 Mbit/s	DL: >500 Mbit/s UL: > 250 Mbit/s	DL: >1.0 Gbit/s UL: >0.5 Gbit/s
Area traffic capacity (Broadband)	DL: >75 Mbit/s/km2 UL: >37 Mbit/s/km2	DL: >750 Mbit/s/km2 UL: >370 Mbit/s/km2	DL: >1.5 Gbit/s/km2 UL: >0.75 Gbit/s/km2
Experienced data rate (NB-IoT)	DL: >2 Kbit/s UL: >10 Kbit/s	DL: >20 Kbit/s UL: >100 Kbit/s	DL: >40 Kbit/s UL: >200 Kbit/s
Area traffic capacity (NB-IoT)	DL: >8 Kbit/s UL: >40 Kbit/s	DL: >80 Kbit/s UL: >400 Kbit/s	DL: >160Kbit/s/km2 UL: >800Kbit/s/km2

[1] User demand that is not satisfied

[2] Used satellite resources such as power, bandwidth, etc

[3] Allocation of satellite resources such as power, spectrum, beampattern given a change in the demand

[4] Increase in successful connections

[5] Reduction with respect to today manual intervention

[6] Gain with respect to 2020 wireless area capacity

[7] % of total operation time

(Table copied from [Networld2020-SRIA] - Networld 2020 (SNS) Strategic and Research Innovation Agenda (SRIA) 2021 – 2027, see: https://bscw.5g-ppp.eu/pub/bscw.cgi/d367342/Networld2020%20SRIA%202020%20Final%20Version%202.2%20.pdf)

The optical community is proposing the following key performance indicators

	Target KPI	Current 2020	Short-term Evo ~2025	Mid-term Evo ~2028	Long-term Evo ~2030
Metro/Core	Spectrum[1]	5THz	15THz	30THz	50THz
	Port speed[2]	400Gb/s	1.6Tb/s	3.2Tb/s	6.4Tb/s
	Bandwidth[3]	<75GHz	<300GHz	<600GHz	<1200GHz
	Line capacity[4]	25Tb/s	200Tb/s	600Tb/s	1.5Pb/s
	Node capacity[5]	150Tb/s	1.2Pb/s	3.6Pb/s	9Pb/s
Access	PON speeds	10Gb/s	50Gb/s	100Gb/s	>200Gb/s
	User data rate[6] (consumer)	100Mb/s	~1Gb/s	>2.5Gb/s	>5Gb/s
	User data rate[6] (business)	1Gb/s	~10Gb/s	>25Gb/s	>50Gb/s
	Latency[7]	<1ms	<100µs	<10µs	<1µs
	Power consumption[8]	100% (baseline)	40%	30%	20%
	Service provisioning	Hour	Min	Second	Sub-second
	Network operations	Operator-controlled, reactive	Intent-based, proactive	Self-diagnosing	Self-optimizing

[1] 25% CAGR, in line with conservative traffic predictions

[2] Extrapolation of Ethernet roadmap

[3] Using 400G DP-16QAM as baseline

[4] 50% CAGR, in line with internet content provider traffic predictions. Assumes exploitation of frequency and space domain.

[5] Based on degree 4 node with 50% local add/drop

[6] 50% CAGR based on Nielsen's law

[7] Excluding propagation delay

[8] 15% reduction per Gb/s p.a., extrapolated from past transponder data

(Table copied from [Networld2020-SRIA] - Networld 2020 (SNS) Strategic and Research Innovation Agenda (SRIA) 2021 – 2027, see: https://bscw.5g-ppp.eu/pub/bscw.cgi/d367342/Networld2020%20SRIA%202020%20Final%20Version%202.2%20.pdf)

With respect to the system architecture and networking the following metrics are proposed:

- Runtime Service Scheduling efficiency increase compared to overprovisioning (for a service requiring 99.999% or higher success rates and under typical traffic arrival conditions)

Short term	Medium term	Long term
2x in single tenant environments	10x in single tenant	At least 10x in multitenant environments

(Table copied from [Networld2020-SRIA] - Networld 2020 (SNS) Strategic and Research Innovation Agenda (SRIA) 2021 – 2027, see: https://bscw.5g-ppp.eu/pub/bscw.cgi/d367342/Networld2020%20SRIA%202020%20Final%20Version%202.2%20.pdf)

This includes aspects as path stretch ((ratio between the average control plane path and the average physical node distance) and resource overhead (services being provided by the network resources versus maximum capacity of those resources).

- Time required for runtime conflict resolution when applying resource efficiency methods, that is the increase in multiplexing desired when compared to independent exclusive allocations and the time that is required to settle all the conflicts that may exist.

Short term	Medium term	Long term
2x for multiple concurrent, overlapping allocations	10x for multiple concurrent, overlapping allocations	At least 10x with critical guarantees

(Table copied from [Networld2020-SRIA] - Networld 2020 (SNS) Strategic and Research Innovation Agenda (SRIA) 2021 – 2027, see: https://bscw.5g-ppp.eu/pub/bscw.cgi/d367342/Networld2020%20SRIA%202020%20Final%20Version%202.2%20.pdf)

- In terms of network-resources collection (network garbage collection), in the sense of recovering resources that are not being used anymore, we expect:

Short term	Medium term	Long term
Feasible, additional recovery process off-line	Feasible, running with the resource allocation	Optimal, on resource allocation actions

(Table copied from [Networld2020-SRIA] - Networld 2020 (SNS) Strategic and Research Innovation Agenda (SRIA) 2021 – 2027, see: https://bscw.5g-ppp.eu/pub/bscw.cgi/d367342/Networld2020%20SRIA%202020%20Final%20Version%202.2%20.pdf)

- Features of the pervasive resource control, in terms of autonomic functions.

	Short term	Medium term	Long term
Configuration	Only a minimal initial pre-configuration (only domain name + security association data, e.g. private/public key)	No human intervention	No human intervention across different domains
Scalability	High, large number of nodes	Very High, any number of nodes, densities	Very High, any number of nodes, densities and complexity
Bootstrapping	Reduced time to 70%	Reduced time to 40%	Reduced time to 10%
Convergence time of the control plane	Time reduced to 70%	Time reduced to 40%	Time reduced to 10%
Signalling overhead in reconfiguration	Reduced to 90%	Reduced to 75%	Reduced to 75% in multitenant environments

(Table copied from [Networld2020-SRIA] - Networld 2020 (SNS) Strategic and Research Innovation Agenda (SRIA) 2021 – 2027, see: https://bscw.5g-ppp.eu/pub/bscw.cgi/d367342/Networld2020%20SRIA%202020%20Final%20Version%202.2%20.pdf)

- In terms of network-suitable AI, it is expected:

Short term	Medium term	Long term
Adaptation of current centric-implementation AI models	Fully distributed AI algorithms at the network	distributed AI supporting and serving several models at the same time

(Table copied from [Networld2020-SRIA] - Networld 2020 (SNS) Strategic and Research Innovation Agenda (SRIA) 2021 – 2027, see: https://bscw.5g-ppp.eu/pub/bscw.cgi/d367342/Networld2020%20SRIA%202020%20Final%20Version%202.2%20.pdf)

In security domain, being a mandatory condition for numerous objectives, security is de facto a pre-requisite for the ongoing Digitalization of our societies. Building trust is combination of awareness, understanding and obviously provision of the right solutions with the right level of security. The ambitious objectives listed below aims at being representative of this combination:

- Towards access to real time Cyber Threat Intelligence information (attacks/threats and vulnerabilities), risk Analysis tools and Services enabling 100% of awareness and level-based appropriate protection counter-measure deployment.

Shor term	Medium term	Long term
Federated, consolidated, common basis across CERTs (CSIRT network, NIS directive application)	CTI platforms (including openCTI) and tools for State-of-The-Art sanitization	100% of qualified threats knowledge and appropriate counter measures made accessible

(Table copied from [Networld2020-SRIA] - Networld 2020 (SNS) Strategic and Research Innovation Agenda (SRIA) 2021 – 2027, see: https://bscw.5g-ppp.eu/pub/bscw.cgi/d367342/Networld2020%20SRIA%202020%20Final%20Version%202.2%20.pdf)

- Trust in ICT infrastructure through systematic Exposure of cybersecurity levels 100% compliant with European-legal basis (certification, Security Service Level attributes, GDPR/EU strategy for Data,...)

Short term	Medium term	Long term
5G systems & services certification frameworks, Basic security level exposure with generic security attributes defined	Methodologies and tools for composition and time evolution of certified perimeters (systems & services)	Evolutive approach for data and disruptive technologies

(Table copied from [Networld2020-SRIA] - Networld 2020 (SNS) Strategic and Research Innovation Agenda (SRIA) 2021 – 2027, see: https://bscw.5g-ppp.eu/pub/bscw.cgi/d367342/Networld2020%20SRIA%202020%20Final%20Version%202.2%20.pdf)

- Compliance with highly critical applications and essential services requirements leading to sovereign solutions able to provide 100% availability of services for verticals

Short term	Medium term	Long term
Local, private implementation for limited set of verticals	End-to-End hybrid implementation for most of verticals	High grade support with technology, system and solution independence

(Table copied from [Networld2020-SRIA] - Networld 2020 (SNS) Strategic and Research Innovation Agenda (SRIA) 2021 – 2027, see: https://bscw.5g-ppp.eu/pub/bscw.cgi/d367342/Networld2020%20SRIA%202020%20Final%20Version%202.2%20.pdf)

- Improve attack detection & response mean time of Cybersecurity incidents including zero % unprotected data leakage

Short term	Medium term	Long term
Benchmark strategy including data set and models	Monitoring and attack detection EU-wide strategy	Data protection strategy with response time and robustness outperforming attackers capabilities

(Table copied from [Networld2020-SRIA] - Networld 2020 (SNS) Strategic and Research Innovation Agenda (SRIA) 2021 – 2027, see: https://bscw.5g-ppp.eu/pub/bscw.cgi/d367342/Networld2020%20SRIA%202020%20Final%20Version%202.2%20.pdf)

Annex IV   Siemens White Paper "5G communication networks: Vertical industry requirements"

In [Siemens2016], several 5G requirements were derived by Siemens based on their studies on vertical application domains, such as Smart City, Smart Mobility, Smart Manufacturing, Smart Energy and Smart Building.

Table 26 shows a consolidated view of the 5G requirements, while **Table 27** provides more details on the 5G requirements coming from verticals.

Table 26: 5G promises vs. Vertical requirements, copied from [Siemens2016]

Category	Requirement	Explicit 5G promises (according to [1], Figure 2)	Consolidated requirements from verticals - Siemens view
Industry-grade Service Quality	Realtime capability – Latency	5 ms (e2e)	1 ms (local) 5 ms (long distance)
	Realtime capability —Jitter	-	1us (local)
	Bandwidth	Peak data 10 Gbps Mobile data volume 10 TB/s/km² Number of devices: 1 mio/km²	kbps … 10Gbps
	Time period of information loss during failures	-	none (seamless failover)
	Availability/coverage	-	ubiquitous
	Range (distance between communication neighbors)	-	0,1 m … 200 km
	Reliability (minimum uptime per year [%])	99,999%	99,9999%
	Mobility	500km/h	500km/h
	Outdoor terminal location accuracy	<1m	0,1 m
	Multi-tenant support	yes (Network Slices)	yes
Operation and maintenance	Non-standard operating conditions	Energy consumption reduced by factor 10	• Battery powered devices with >10years lifetime • Harsh environments (weather, vibrations, heat, dust, hazardous gases, etc.)
	Ease of use	-	• Communication services approach • Plug and play device (sensor, actuator, controller) integration
	SLA Tooling	-	Service Level Agreement (SLA) monitoring and management tools for provider and consumer
	Service deployment time (time between service request and service realization)	90 min	hours
	Private 5G infrastructures	-	yes
Non-technical	Scalability: Number of devices per km²	10^6	10^5
	Globally harmonized definition of Service Qualities	-	yes
	Technology availability	-	>20 years
	Globally simplified certification of ICT components	-	Yes
Assured Guarantees		-	mandatory

Table 27: 5G promises vs. Vertical requirements (details), copied from [Siemens2016]

Category	Requirement	Explicit 5G promises (according to [1], Figure 2)	Siemens demand	Smart City	Smart Mobility	Smart Manufacturing		Smart Energy			Smart Building
						Process	Discrete	Low Voltage	Medium Voltage	High Voltage	
Industry-grade Service Quality	Realtime capability – Latency	5 ms (e2e)	1 ms (local) 5 ms (long distance)	-	1ms (local) 10 ms (long distance)	20ms (local) 1s (long distance)	1ms (local) 20ms (long distance)	-	25ms	5ms (long distance)	100ms
	Realtime capability — Jitter	-	1us (local)	-	-	20ms	1us	-	25ms	1ms	-
	Bandwidth	Peak data 10 Gbps Mobile data volume 10 TB/s/km^2 Number of devices: 1 mio/km^2	kbps ... 10Gbps	kbps (sensors) ... Mbps (video supervision) ... 10 Gbps (data centers)	10 Mbps ... 1 Gbps	100 kbit/s (automaton stream) ... 100 Mbps (remote access, video supervision)	100 kbit/s (automaton stream) ... 100 Mbps (remote access, video supervision)	1 kbps per subscriber	5 Mbps per secondary substation	1Gbps along power lines	100 kbit/s (automaton stream) ... 100 Mbps (remote access, video supervision)
	Time period of information loss during failures	-	none (seamless failover)	1s	100 ms	100 ms	none (seamless failover)	minutes	25ms	none (seamless failover)	100 ms
	Availability/coverage	-	Ubiquitous	City-level	Ubiquitous	Industrial Plant Areas	Industrial Plant Areas	Ubiquitous	Ubiquitous	Ubiquitous	City-level
	Range (distance between communication neighbors)	-	0,1 m ... 200 km	10 km	1 km (cars) ... 10 km (trains)	0,1m ... 10 km	0,1 m ... 100 m	10 km	20 km	200 km	100m
	Reliability (minimum uptime per year [%])	99,999%	100%	99,9%	100%	100%	100%	98%	99,9%	100%	99,9%
	Mobility	500km/h	500km/h	100km/h	500km/h	50km/h	50km/h	5km/h	-	-	5km/h
	Outdoor terminal location accuracy	<1m	0,1 m	1 m	0,1 m	0,1 m	0,1 m	10 m	10 m	-	0,1 m
	Multi-tenant support	yes (Network Slices)	yes								
Operation and maintenance	Non-standard operating conditions	Energy consumption reduced by factor 10	• Battery powered devices with >10years lifetime • Harsh environments (weather, vibrations, heat, dust, hazardous gases, etc.)								
	Ease of use	-	• Communication Services approach • Plug and Play Device (Sensor, Actuator, Controller) integration								
	SLA Tooling	-	Service Level Agreement (SLA) monitoring and management tools for provider and consumer								
	Service deployment time (time between service request and service realization)	90 min	hours								
	private 5G infrastructures	-	yes	-	yes	yes	yes	-	optional	yes	optional
Non-technical	Scalability: Number of devices per km^2	10^6	10^5	10^5	10^4	10^5 (high density of devices)	10^5 (high density of devices)	10^4	10^3	10^3	10^5
	Globally harmonized definition of Service Qualities	-	yes	-	yes	yes (for long distance)	yes (for long distance)	-	yes	yes	-
	Technology availability	-	>20 years								
	Globally simplified certification of ICT components	-	Yes								
Assured Guarantees		-	Mandatory	Relaxed	Mandatory	Mandatory	Mandatory	Relaxed	Mandatory	Mandatory	Relaxed

Contributors

The document was written by several participants of the AIOTI WG Standardisation.

Editor:

Georgios Karagiannis, Huawei

Reviewer:

Damir Filipovic, AIOTI, Secretary General

Main Contributors:

Agnia Codreanu (BEIA)

Antonio Kung (Trialog)

Antonio Lalaguna (ACISA)

Antonio Skarmeta (University of Murcia)

Arne J. Berre (Sintef)

Artur Krukowski (RFSAT)

Asbjørn Hovstø (Hafenstrom)

Christian Kloch (Force Technology)

Christophe Gossard (John Deere)

Damir Filipovic (AIOTI Secretary General)

Erwin Schoitsch (Austrian Institute of Technology - AIT)

Flemming Sveen (Hafenstrom)

George Suciu (BEIA)

Georgios Karagiannis (Huawei)

Giacomo Tavola (Politecnico di Milano)

Gianmarco Baldini (EC JRC)

Izzet Saglam (Turkcell)

Jaume Segura (Universitat de València)

Joao Peixoto (Ubiwhere)

Jose Luis Hernandez (EC JRC)

Kevin McDonnell (Huawei)

Konstantinos Loupos (INLECOM)

Konstantinos Ntafloukas (INLECOM)

Krzysztof Piotrowski (IHP-Microelectronics)

Lavinia Ana Petrache (BEIA)

Marco Carugi (Huawei)

Mari-Anais Sachian (BEIA)

Maria Niculae (BEIA)

Mihai Penica (University of Limerick)

Natalie Samovich (Enercoutim)

Nemanja Mišić (DunavNET)

Nikos Giannakakos (UniSystems)

Ranga Rao Venkatesha Prasad (Technical University Delft)

Rui Aguiar (University of Aveiro)

Rute Sofia (fortiss)

Sean McGrath (University of Limerick)

Srđan Krčo (DunavNET)

Thomas Klein (IBM)

Toon Norp (TNO)

Valentina Peniche (AIOTI)

Vasileios Karagiannis (Austrian Institute of Technology - AIT)

Zbigniew Kopertowski (Orange)

Acknowledgements

About AIOTI

AIOTI is the multi-stakeholder platform for stimulating AI, IoT and Edge Continuum Innovation in Europe, bringing together small and large companies, academia, researchers, policy makers, end-users and representatives of society in an end-to-end approach. We strive to leverage, share and promote best practices in the AI, IoT and Edge Continuum ecosystems, be a one-stop point of information to our members while proactively addressing key issues and roadblocks for economic growth, acceptance and adoption of the AI, IoT and Edge Continuum Innovation in society. AIOTI contributions goes beyond technology and addresses horizontal elements across application domains, such as matchmaking and stimulating cooperation by creating joint research roadmaps, defining policies and driving convergence of standards and interoperability.