6G LEADER

LEARNING-DRIVEN AND EVOLVED RADIO FOR 6G COMMUNICATION SYSTEMS

D2.1 Use Case Analysis, KPIs and Requirements to RAN Architecture Design

30/09/2025

Funded by the European Union

6GSNS

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the Smart Networks and Services Joint Undertaking. Neither the European Union nor the granting authority can be held responsible for them.

Grant Agreement No.	101192080
Project Acronym/ Name	6G-LEADER: LEArning-Driven and Evolved Radio for 6G Communication Systems
Topic	HORIZON-JU-SNS-2024-STREAM-B-01-02
Type of action	HORIZON-JU-RIA
Service	SNS
Duration	36 months (starting date 1 January 2025)
Deliverable title	Use case analysis, KPIs and requirements to RAN architecture design.
Deliverable number	D2.1
Deliverable version	v1.0
Contractual delivery date	30 / 09 / 2025
Actual date of delivery	30 / 09 / 2025
Nature of deliverable	Report
Dissemination level	Public
Work Package	WP2 (Task 2.1&Task 2.2)
Deliverable lead	Accelleran ACC
Author(s)	German Castellanos, Stephen Parker (ACC), Constantinos Psomas, Themistoklis Charalambous (UCY), Cristina Costa, Filippo Cugini (CNIT), Marios Kountouris (UGR), Federico Mungari, Enrique Marín (ATOS), Farhad Mehran (DICAT), Rolando Guerra-Gomez (TEF), Petros Maratos, Grigorios Kakkavas, Maria Diamanti, Anastasios Zafeiropoulos, Symeon Papavassiliou (ICCS) and Haya Al Kassir (FDI), Tugce Kobal (NOKIA), Matthew Andrews (NOKIA). Oriol Font (SRS).
Abstract	D2.1 defines a top-down method to translate 6G-LEADER's vision into an O-RAN-based architecture. It maps SDGs, to innovation pillars, to objectives, to high-level use cases, to KPIs and E2E requirements, and shows how AI-driven communication techniques and reconfigurable RF components support those use cases. The deliverable sets the KPI/requirements framework, and provides traceable inputs for architecture, PoC design, and validation in later WPs.
Keywords	6G, RAN architecture, SDGs, innovation pillars, use cases, KPIs, KVIs, AI/ML, PHY/MAC, semantic communications, AirComp, Wireless for AI, RIS, fluid antennas, reconfigurable RF, conflict management, edge computing, PoC validation.

Disclaimer

Copyright notice

Contributors

Organisation	Name	Contribution
ACC	German Castellanos and Stephen Parker.	Main editor, Chapter 1, 2, 3 and 7 contributions. Conclusions, future work.
ATOS	Enrique Marin and Federico Mungari.	O-RAN architecture and conflict management challenges description in Section 5.
AALTO	Risto Wichman and Xichun Chen.	Chapter 4 contributions.
CNIT	Cristina Costa and Filippo Cugini.	Chapter 3 (UCG2 - XR & UAV seamless interaction) and Chapter 4 contributions (Data collection and lifecycle, Energy-efficient semantics-aware schemes).
DICAT	Farhad Mehran.	Chapter 3 (use case overview) and Chapter 5 (O-RAN Conflict Management)
FDI	Haya Al Kassir.	Chapter 3,4,5, and 6 contributions.
ICCS	Grigorios Kakkavas, Petros Maratos, Maria Diamanti, Anastasios Zafeiropoulos and Symeon Papavassiliou.	Chapter 4 contributions.
MB	Andreas Benzin and Fabian Göttsch.	RIS for FR3 hybrid beamforming and system-level considerations (coverage, cost/complexity).
NOKIA	Tugce Kobal and Matthew Andrews.	Chapter 4 and 5 contributions.
SRS	Oriol Font.	SRS CU/DU platform used across PoCs; O-RAN extensions and SMO interface mapping.
SRUK	Mutasem Hamdan.	Conflict management in O-RAN.
UC3M	Ana Garcia and Javier Otero.	FA prototype and support to FR1/FR3 coexistence evaluations.
UCY	Constantinos Psomas and Themistoklis Charalambous.	Chapter 1, Chapter 4 and Chapter 5 contributions.
UGR	Marios Kountouris.	Chapter 4 contributions.
UPC	Judit Cerda.	ML-aided channel learning/prediction.

Peer Reviewers

Name	Organisation
Konstantinos Ntougias	UCY
Matthew Andrews	NOKIA

Document Revision History

Version	Date	Owner	Author(s)	Comments
0.1	12/03/25	ACC	German Castellanos, Stephen Parker (ACC), Constantinos Psomas (UCY).	Document creation and table the contents agreed.
0.2	15/04/25	ACC	Haya Al Kassir (FDI), Rolando Guerra Gomez (TEF), Marios Kountouris (UGR), Cristina Costa (CNIT).	Initial partners contributions for Chapters 2, 3 and 4.
0.3	14/05/25	ACC	Filippo Cugini (CNIT), Haya Al Kassir (FDI), Tugce Kobal (NOKIA), Marios Kountouris (UGR).	Initial partners contributions for Chapters 4, 5 and 6.
0.4	21/05/25	ACC	German Castellanos (ACC).	Merging several contributions and format initial revision.
0.5	01/09/25	ACC	German Castellanos (ACC).	Final review draft.
0.9	25/09/25	ACC	German Castellanos (ACC).	Modification based on internal review process.
1.0	30/09/25	F6S	Mateusz Kowacki (F6S).	Final Version.

TABLE OF CONTENTS

LIST OF FIGURES

List of Tables

ABBREVIATIONS AND ACRONYMS

3GPP	3rd Generation Partnership Project
AA	Access Aware
AI	Artificial Intelligence
AirComp	Over-the-air Computing
AN	Anchor Node
AoI	Age of Information
AP	Access Point
AR	Autoregressive
BS	Base Station
CDM	Conflict Detection and Management
CIO	Cell Individual Offset
CoUD	Cost of Update Delay
CS	Compressed sensing
CSI	Channel state information
CU	Central Unit
D2D	Device to Device
dApp	Distributed Application
DNN	Deep Neural Network
DQN	Deep Q-Network
DRL	Deep Reinforcement Learning
DSA	Dynamic Spectrum Access
DU	Distributed Unit
E2E	End to End
eGaIn	Eutectic Alloy of Gallium and Indium
eHE	eHealth and Emergency
ETSI	European Telecommunications Standards Institute
FA	Fluid Antenna
FAMA	Fluid Antenna Multiple Access
FL	Federated Learning
FLMR	Federated Machine Reasoning
GAN	Generative Adversarial Network
GenAI	Generative Artificial Intelligence
GPR	Gaussian Process Regression
GRU	Gated Recurrent Unit
HDTGA	Hierarchical Decision Transformer with Goal Awareness
IBN	Intend-Based Networking
ILM	Intent-Lifecycle-Management
IoRT	Internet of Robotic Things
IoT	Internet of Things
ISAC	Integrated Sensing and Communication
ITU	International Telecommunication Union
KPI	Key Performance Indicator
KSBS	Kalai-Smorodinsky Bargaining Solution
KVI	Key Value Indicator
LCM	Life Cycle Management
LLM	Large Language Model
LMMSE	Linear Minimum Mean Square Error
LS	Least Squares
LSTM	Long Short-Term Memory
MAC	Medium Access Control
MAF	Methodological Assessment Framework
MARL	Multi-Agent RL
MDP	Markov Decision Process
MIMO	Multiple Input Multiple Output
ML	Machine Learning
mMTC	massive Machine Type Communication
MNO	Mobile Network Operator
MRC	Maximum Ratio Combining
MSE	Mean Squared Error
NF	Network Function
NFED-RIS	Near-Field Fed Reflective Intelligent Surface
nGRG	next Generation Research Group
NIF	Network Intelligent Function
NIO	Network Intelligent Orchestrator
NIS	Network Intelligent Services
NOMA	Non-Orthogonal Multiple Access
NT	Network Tomography

NWDAF	Network Data Analytics Function
OFH	Open Front Haul
OMA	Orthogonal Multiple Access
O-RAN	Open Radio Access Network
OWL	Ontology Web Language
PHY	Physical (layer)
PoC	Proof of Concept
PPDR	Public Protection and Disaster Relief
PPO	Proximal Policy Optimization
QoE	Quality of Experience
QoS	Quality of Service
RAG	Retrieval Augmented Generation
RAN	Radio Access Network
rApp	RAN Application
RB	Resource Block
RDF	Resource Description Framework
RF	Radio Frequency
RIC	RAN Intelligent Controller
RIS	Reconfigurable Intelligent Surface
RL	Reinforcement Learning
RNN	Recurrent Neural Network
RSMA	Rate-Splitting Multiple Access
RT	Real-Time
RU	Radio Unit
SDG	Sustainable Development Goal

SEBS	Shannon Entropy Bargaining Solution
SIC	Successive Interference Cancellation
SIR	Signal to Interference Ratio
SMO	Service and Management Orchestration
SN	Server Node
SNS JU	Smart Networks and Services Joint Undertaking
SOTA	State of the Art
SQL	Structured Query Language
S-RAN	Semantic Aware Radio Access Network
UC	Use Case
UCG	Use Case Group
UL	Uplink
URLLC	Ultra-Reliable Low-Latency Communications
VAE	Variational autoencoder
VoI	Value of Information
VoIU	Value of Information of Update
vRAN	Virtualized RAN
WDA	Wireless Data Aggregation
WNBS	Weighted Nash Bargaining Solution
WP	Work Package
XAI	Explainable AI
xApp	Extensible Application

Executive Summary

This deliverable defines how **6G-LEADER** turns its vision into a testable architecture based on an extensive project requirement analysis. It establishes a top-down method that starts from societal drivers and Sustainable Development Goals (SDG) alignment and flows through innovation pillars, project objectives and high-level use cases to measurable Key Performance Indicators (KPI) and end-to-end requirements. The document then connects these elements into clear technological elements that provide the means to implement the objectives of the project while drawing an initial O-RAN-based RAN architecture and to five Proof-of-Concepts (PoC) that will validate the approach on real testbeds. In doing so, this deliverable D2.1 provides a clear technological framework that establish the project architecture.

The methodology is straight forward: SDGs define why act; innovation pillars state where to act while objectives explain what must be achieved. Use cases depict where the technology is exercised; KPIs quantify its success while translating them into implementable targets; finally, the architecture and PoCs operationalise all. This methodology allows later updates.

On content, the deliverable advances three areas. First, it maps SDGs to project impact, showing clear contributions to health, education, energy efficiency, resilient infrastructure and partnerships. Second, it defines seven innovation pillars that capture the project's technology scope: AI/ML-driven PHY; multiple access and Wireless-for-AI (incl. AirComp); highly reconfigurable RF such as fluid antennas and RIS, FR1/FR3 coexistence; semantics-empowered communications and AI/ML-driven techniques for goal-oriented semantic networking; real time RAN control; and conflict management across x/r/dApps. Third, it derives objectives and a KPI catalogue that quantify expected gains, such as halving E2E latency, improving spectral efficiency, reducing energy and EMF, and enabling sub-10 ms control loops.

The requirements mapping links each objective to KPIs and to the PoCs that will exercise them. Five PoCs cover XR–UAV real-time interaction with semantic video, FR3 eMBB with hybrid/RIS beamforming, conflict-aware RIC control for energy efficiency and traffic steering, AirComp-enabled Wireless-for-AI with semantic offloading, and AI-aided multiple access with fluid antennas.

The deliverable also outlines how enabling techniques translate into system-level benefits. AirComp reduces aggregation latency and radio overhead for distributed learning and control. Semantic communications cut non-useful traffic and improve timeliness. Predictive access and scheduling increase spectral and energy efficiency. RT RIC logic and a conflict manager turn these gains into stable operation under realistic load. On the RF side, RIS and fluid antennas provide additional performance which is exposed through controllable interfaces to the RIC.

D2.1 is positioned at the front of the technical pipeline. WP3–WP6 use its objectives, KPIs and requirements to focus research on PHY/MAC, RIC and reconfigurable RF; WP7 plans and executes PoC validation against the mapped KPIs, with results feeding back into requirements and targets. Looking ahead, WP2 will extend the architecture in Task 2.4. With this, D2.1 provides a stable baseline and a practical path to demonstrate measurable progress on real platforms.

1 Introduction

This deliverable ***D2.1 Use case analysis, KPIs and requirements to RAN architecture design***, sets out how **6G-LEADER** translates its vision into a concrete RAN architecture based on new components envisioned for 6G networks. It introduces a top-down methodology that starts from societal drivers (SDGs) and flows through innovation pillars, project objectives and high-level use cases to measurable KPIs and end-to-end requirements. Via mapping the requirements into use cases groups, D2.1 described how the envisioned components are supporting these uses case to define a low-level mapping. This brings a detailed description of the components such advanced AI-driven communication techniques and reconfigurable RAN components, that are the basis for the **6G-LEADER** RAN architecture.

The deliverable fulfils three main objectives. First, it defines the framework for requirements engineering in **6G-LEADER**, on how use cases are specified, which KPIs matter, and how those KPIs translate into E2E performance targets. Second, it provides the traceability from SDGs and objectives to architecture choices and Proof-of-Concepts (PoCs) plans, so that every technical decision can be justified and measured. Third, it prepares the ground for implementation by consolidating the initial architecture view and the PoC mapping that will be used for testing and validation in subsequent work packages.

The following sections describe in detail the objectives of the deliverable and how it is structured.

1.1 Scope and objectives of D2.1

Deliverable D2.1 presents a high-level description of representative 6G use cases, which serves as the foundation for defining system-level requirements and establishing the design process of the **6G-LEADER** O-RAN-based architectural framework.

The deliverable encompasses the outputs of the first two tasks of Work Package (WP) 2, Tasks 2.1 and 2.2, based on two main activities: (i) the identification and analysis of emerging technologies and signal processing techniques relevant to the 6G physical (PHY) layer and O-RAN, and (ii) the definition of use cases, PoC requirements and **6G-LEADER**'s initial architecture. As part of Task 2.1, deliverable D2.1 includes a comprehensive technology radar covering recent advancements in areas such as reconfigurable Radio Frequency (RF) components, semantic communication, over-the-air computation and AI/ML-enhanced PHY layer techniques. These findings have a direct impact on architectural decisions within the **6G-LEADER** framework. From Task 2.2, it integrates detailed requirements extracted from the defined use cases and PoC scenarios. These are aligned with the **6G-LEADER**'s objectives and sustainability principles, laying the foundation for the subsequent project phases and implementation strategies.

The deliverable adopts a top-down methodology to ensure that the **6G-LEADER** architecture aligns with societal needs and long-term impact goals. The process begins with mapping the project's envisioned contributions to the UN Sustainable Development Goals (SDGs), providing the basis for social and environmental impact [1]. Building on this, the project's strategic objectives and innovation pillars define the main technological directions. A set of high-level use cases is then derived to guide functional priorities and real-world relevance. Each use case is translated into measurable Key Performance Indicators (KPIs), Key Value Indicators (KVIs) and associated technical requirements, ensuring that the system design remains focused on specific performance targets. Finally, these are integrated into the architectural design and PoCs, creating a validation loop that connects high-level goals (i.e. SDGs) with low-level implementation, enabling iterative refinement throughout the **6G-LEADER** lifecycle.

In summary, the main objectives of this deliverable are to:

- Define a top-down methodology that links UN SDGs, innovation pillars, technical objectives, KPIs and PoCs.
- Map **6G-LEADER**'s research and architectural vision into concrete use cases and system-level requirements.
- Identify and structure the innovation areas critical for enabling intelligent, sustainable and efficient 6G networks.
- Establish a comprehensive KPI framework for tracking project progress and validating the performance of core technologies.
- Provide the foundation for system design, validation and PoC alignment in future WPs.

1.2 Document structure and relation to other WPs

This deliverable D2.1, part of the WP2, sits at the front of the project's technical pipeline and feeds the downstream work packages with traceable, testable inputs. From WP2's SDG-to-requirements mapping and KPI framework, WP3–WP5 derive concrete research targets: WP3 advances AI/ML-enhanced PHY/MAC enablers and AirComp; WP4 designs and prototypes highly reconfigurable RF components (e.g., RIS, fluid antennas) and spectrum-coexistence strategies; WP5 develops goal oriented and semantic empowered communications. WP6 consolidates these outputs into a coherent **6G-LEADER** RAN architecture supported by AI/ML and semantic extension to the O-RAN based architecture focused on RAN control, xApp/rApp/dApp logic and system-level optimisation. WP7 then uses the same traceability chain (Objectives → KPIs → Requirements) to schedule, implement, and evaluate the project's Proof-of-Concepts on the selected testbeds, ensuring that validation aligns with the KPI targets defined in this document. In this way, the D2.1 feeds from all the project WPs to create a clear use case analysis framework until the architecture design.

The methodology Chapter 2 explains the top-down process used in the project, from SDGs and innovation pillars through to use cases, KPIs and end-to-end requirements. It details how each step produces artefacts that can be verified later (e.g., KPI targets, test conditions) and how these artefacts are versioned to support iterative refinement with WP3–WP6.

Then, Chapter 3, translates the high-level vision into concrete needs per use case. It documents functional and performance requirements, shows their linkage to KPIs and objectives, and provides the traceability matrix that WP6 uses for architectural decisions and WP7 uses for PoC acceptance criteria.

Chapter 4 dives on advancements in communication techniques summarising the enabling mechanisms developed (or adopted) in the project—AI/ML-driven PHY/MAC, semantics-aware transmission, wireless computation (e.g., AirComp), traffic steering, and reliability/latency optimisations. It states the expected KPI impact per mechanism and outlines test hooks for later validation. Chapter 5 discuss the reconfigurable components covering the RF side: fluid antennas, RIS-assisted beamforming, and FR1/FR3 coexistence enablers. It defines their roles in energy efficiency, EMF reduction, and spectral efficiency, and specifies the interfaces and measurements required so WP6 can embed them and WP7 can evaluate them consistently.

Next, the architecture chapter 6 presents the **6G-LEADER** RAN view that integrates these capabilities. It describes functional blocks (e.g., near-RT RIC, data/AI pipelines, RU/DU/CU splits), control loops (xApp/rApp/dApp), and north/southbound interfaces, and shows how the design satisfies the requirements and KPI targets traced from earlier chapters.

Finally, Chapter 7 closes the deliverable with the status of requirements coverage, the readiness of architectural elements for integration, and the handover to WP6/WP7 for implementation and validation, ensuring the project remains aligned with its objectives and KPI commitments.

2 Methodology

The development of a consistent and measurable methodology is essential to ensure that the **6G-LEADER** project's innovations are not only technically sound but also aligned with broader societal and environmental objectives. Chapter 2 introduces the approach adopted by the project to systematically identify, categorise, and map KPIs, KVIs, and high-level Use Cases (UCs) to the project's innovation pillars and architecture. This mapping process serves as a backbone for validating technological progress and demonstrating impact through real-world PoCs.

The methodology is based on the strategic frameworks defined by the Smart Networks and Services Joint Undertaking (SNS JU) [2], and it is informed by state-of-the-art practices from other Stream B and D projects. By leveraging the structure proposed in the SNS KPI and KVI white paper [3], [4] and building upon the experiences from key initiatives such as FIDAL, Hexa-X, TrialsNet [5], [6], [7], and others, **6G-LEADER** ensures coherence, comparability, and alignment with European 6G research goals.

The following sections outline the landscape of KPI/KVI mapping methodologies across JU-SNS projects, followed by a detailed description of the mapping process implemented in **6G-LEADER**.

2.1 Mapping to other JU-SNS projects

The current State of the Art (SOTA) regarding the definition and mapping of KPIs, KVIs, and Use Cases within the SNS-JU projects has evolved significantly, guided by foundational frameworks and enriched by a diversity of project-specific contributions. At the heart of the harmonization efforts stands the SNS Test, Measurement, and KPIs Validation Working Group, whose "*6G KPIs – Definitions and Target Values*" white paper establishes a comprehensive foundation for KPI categorization and validation across projects [4]. The SNS paper classifies KPIs into the following clearly defined families: Data Rate, Latency, Reliability, Mobility, Sensing, EMF, AI, Positioning, Energy Efficiency, Coverage, Compute, and Other KPIs as described in Table 2.1. This structure not only aligns KPI definitions with international standards like the 3rd Generation Partnership Project (3GPP) and International Telecommunication Union (ITU) but also introduces methodologies to bridge the evaluation of technical performance with KVIs, thus ensuring that technological innovations are matched by societal and environmental impacts [3][4].

Among the most advanced individual projects, *FIDAL* presents a meticulous framework for validating KPIs and KVIs within Public Safety and Media verticals. KPIs in *FIDAL* address metrics such as Application Latency, Positioning Accuracy, High Throughput, and Content Load Times, providing a layered evaluation of network application deployment times and service quality across use cases like Digital Twin for First Responders and XR-assisted services for public safety.

Table 2.1. KPI families from SNS white paper.

KPI Family	Description
Data Rate	Measures user-experienced and peak data throughput
Latency	Measures end-to-end transmission delay and application-level delays
Reliability	Ensures successful data delivery within the required time constraints
Mobility	Assesses network support for user mobility, including handovers
Sensing	Evaluates network capabilities to sense and interpret physical environments
EMF	Addresses electromagnetic field exposure constraints
AI	Evaluates AI integration, such as AI model accuracy and inference latency
Positioning	Measures accuracy and reliability of location-based services
Energy Efficiency	Assesses energy consumption per transmitted bit or service session
Coverage	Evaluates geographical and service availability coverage
Compute	Measures edge/cloud computing performance relevant to network services
Other KPIs	Captures additional indicators like security, trust, and resilience

Moreover, *FIDAL* expands into the domain of KVIs by systematically identifying societal impacts, such as improving safety, cultural access, and environmental sustainability through measured indicators like stakeholder perception of safety and reduced energy usage [5], [8], [9]. Similarly, *Hexa-X-II* offers one of the most complete evolutions in KPI and KVI methodologies. Building upon the legacy of Hexa-X, the project defines six major use case families, including Collaborative Mobile Robots, Physical Awareness, and Immersive Experiences. *Hexa-X-II* pioneers the integration of AI-native KPIs, novel metrics for joint sensing and communication, and methods for environmental sustainability evaluation. The project not only defines KPIs like latency, positioning accuracy, and AI inference delay but systematically correlates them to societal goals by introducing cross-domain KVIs addressing economic growth, security, privacy, and environmental responsibility [6], [10], [11], [12].

Expanding further into JU-SNS Stream D projects, *TrialsNet* represents a flagship effort in large-scale field experimentation. *TrialsNet*'s methodology encompasses 13 use cases, from Smart Crowd Monitoring and Smart Ambulance to City Parks in the Metaverse. The project iteratively harmonizes KPIs across these use cases, drawing from the European Telecommunications Standards Institute (ETSI) and the 3GPP references while remaining open to future evolutions of KPI methodologies. KPIs such as uplink throughput, end-to-end latency, and service reliability are complemented by emergent KVIs aimed on societal resilience, inclusiveness, and trust [11], [13], [14], [15]. In addition, *TARGET-X* focuses on the cross-industrial deployment of 5G and beyond technologies, covering sectors like Manufacturing, Automotive, Energy, and Construction. The project establishes a Methodological Assessment Framework (MAF) for systematically deriving KPIs and KVIs, associating techno-economic benefits with specific use cases. *TARGET-X* identifies uniform use case descriptions and defines 17 KPIs and 10 KVIs to ensure a balanced evaluation of technological performance and societal impact, with exemplary equations for metrics like electric power consumption and water usage [16], [17]. Moreover, *IMAGINE-B5G* introduces a highly structured approach to platform KPIs, leveraging OpenTAP testing methodologies and focusing on verticals such as eHealth, Education, Industry 4.0, and Media [18], [19]. Detailed KPI families include user-experienced data rate, spectral efficiency, E2E latency, and service availability. Beyond technical metrics, *IMAGINE-B5G* frames KVIs around environmental sustainability, economic growth, innovation, inclusiveness, and health improvement, ensuring a full-circle evaluation of their technological and societal impact [20], [21].

Shifting into JU-SNS Stream B projects, *ORIGAMI* contributes to AI-driven network optimization with a focus on trustworthy 6G infrastructures. Although still emerging, *ORIGAMI* identifies early KPIs and KVIs linked to reliability, trust, energy efficiency, and transparent decision-making frameworks, highlighting their vision for an adaptive and resilient network environment [22]. On the other hand, *PRIVATEER* emphasizes privacy and trust as central pillars for 6G evolution. The project identifies KPIs around data protection, secure connectivity, and resilient network architectures. *PRIVATEER* also targets KVIs linked to the enhancement of human rights, digital sovereignty, and the fostering of user trust in interconnected ecosystems [23]. *Deterministic6G* explores the rigorous domain of deterministic communication for critical industrial and robotic applications. Their KPI set focuses heavily on Ultra-Reliable Low-Latency Communications (URLLC), service availability, and timing precision. The KVIs are oriented toward the enhancement of industrial productivity, safety, and energy-efficient operations [24], [25]. Similarly, *SAFE-6G* addresses safety-driven innovations in 6G networks, particularly related to security, resilience, and public protection. The project defines KPIs for service continuity, fault tolerance, and response time, and maps KVIs towards societal protection goals and the enhancement of critical infrastructure safety [26], [27]. Parallelly, *6GTandem* innovates around digital twin integration with network intelligence. KPIs identified include digital twin accuracy, synchronization latency, and prediction reliability. The project complements these with KVIs related to digital empowerment, data sovereignty, and operational efficiency improvements [28], [29], [30]. Finally, *PREDICT-6G* focuses on deterministic AI-native network infrastructures. The project promotes KPIs regarding predictable latency, distributed AI performance, and proactive fault management, while KVIs emphasize trust in autonomous systems, sustainability through intelligent resource allocation, and the democratization of network capabilities [31], [32].

Table 2.2. Main use cases managed by JU-SNS projects.

Project	Main Use Case Themes
PREDICT-6G	Deterministic communication, Smart Factory, AI predictive services
Deterministic6G	Industrial automation (Manufacturing, Exoskeletons), Edge Computing, Digital Twins
ORIGAMI	AI-driven Network optimization, Trustworthy 6G Infrastructure, Sustainability architecture
Hexa-X-II	Collaborative Mobile Robots, Industrial automation, Digital Twins, XR and Smart Cities
IMAGINE-B5G	Emergency services, eHealth (Remote Proctoring, Smart Ambulance), Industry 4.0, Education, Media
TARGET-X	Manufacturing, Automotive, Energy, Construction (5G for cross-industrial digital transformation)
FIDAL	Public Safety (PPDR), Advanced Media services
TrialsNet	Smart Cities, Metaverse applications, Remote Health Monitoring, Emergency Rescue, Smart Infrastructure

Altogether, the SOTA of KPI and KVI definition within the SNS ecosystem shows a remarkable progression towards not only achieving technical excellence but ensuring that societal, environmental, and economic values are embedded at the core of next-generation communication networks, as presented in the Table 2.2.

2.2 Description of the mapping

The mapping methodology of KPIs in the **6G-LEADER** project follows a structured sequence of steps designed to align technological innovation with socio-economic and environmental priorities. The process begins with the identification of relevant SDGs, which are explicitly correlated to the project's objectives to ensure that technical advancements contribute meaningfully to broader global challenges. These project objectives are then mapped to specific innovation pillars defined in the proposal, such as AI-driven physical layer enhancements, reconfigurable RF components, and semantics-empowered communication strategies. These pillars act as the core thematic axes around which technical development is organised.

Building on the pillars, the project defines a set of high-level use cases, reflecting realistic scenarios including XR and UAV real-time interaction, enhanced mobile broadband experiences, and AI-powered RAN management. Each use case introduces specific performance demands, which are translated into high-level KPIs. These KPIs include, but are not limited to, peak data rates, user density, latency bounds, and energy efficiency metrics. Once the KPIs are defined, the process advances to the derivation of End-to-End (E2E) system requirements. These requirements operationalise the KPIs into quantifiable targets and constraints, which in turn guide the specification and design of the overall **6G-LEADER** network architecture. This ensures consistency between system capabilities and the diverse demands of the envisioned use cases.

The architecture is subsequently validated through a set of targeted PoCs. These PoCs are designed to demonstrate, in controlled yet realistic environments, that the system and components can meet the defined performance targets. Each PoC is defined based on its associated KPIs and E2E requirements, ensuring measurable and outcome-oriented validation. The full definition and planning of these PoCs, along with the associated evaluation methodologies, are described in Deliverable D7.1 [33].

This methodical, traceable approach enables rigorous validation, facilitates continuous feedback and refinement, and ensures that the project outcomes are both technically robust and aligned with the strategic goals of the JU-SNS programme, as described in Figure 2.1.

Figure 2.1. General mapping methodology for 6G-LEADER.

The following chapter expands on this methodology by detailing the mapping of requirements, beginning with a high-level description of project-wide needs and associated use cases. It then presents a granular analysis of how these requirements are decomposed and integrated into the system architecture, concluding with an overview of the traceability mechanisms that ensure consistency from use case to implementation.

3 Mapping of the Requirements

This chapter links the project vision to concrete, testable requirements. It applies the top-down method introduced earlier—starting from SDG alignment and innovation pillars, through project objectives and high-level use cases—to derive measurable KPIs and end-to-end requirements. The result is a traceable chain that explains why each requirement exists, which objective it serves, and how success will be measured in the planned PoCs and testbeds. The first part presents the high-level mapping and the **6G-LEADER** pillars and objectives, then enumerates the KPI set that will be used as acceptance targets for validation.

Building on that, the chapter groups related scenarios into Use Case Groups (UCGs) and shows how the enablers (e.g., AirComp, semantic processing, FR3/RIS/FA, conflict-aware RIC control) translate into KPI impact. The closing tables provide the low-level mapping from KPIs to each UCG and PoC, defining what must be instrumented, which loops (near-RT/non-RT/dApps) are involved, and the expected performance gains (latency, spectral/energy efficiency, EMF). This establishes the inputs WP6 needs for architectural decisions and the criteria WP7 will use for planning and evaluating trials.

3.1 High level mapping

The High-level mapping aims to link the **6G-LEADER** in a top-down manner, with the SDG, the Innovation pillars, the objectives KPI and PoC to define a coherent implementation and validation methodology.

The SDGs [1], [34] were mapped against the contributions of the **6G-LEADER** project by systematically analysing the project's objectives, use cases, and architectural innovations. The mapping prioritises real impacts on society, environment, and industry that can be reasonably expected from the technological solutions that **6G-LEADER** is developing. Goals such as Good Health and Well-being (Goal 3) are addressed through the project's focus on enabling low-latency, ultra-reliable communication for remote healthcare applications, such as remote surgery and real-time health monitoring. Similarly, Quality Education (Goal 4) is supported by the project's advancements in immersive technologies, including virtual and augmented reality experiences that can facilitate digital education through high-fidelity simulations and twinning platforms.

Many goals, such as No Poverty (Goal 1), Zero Hunger (Goal 2), Gender Equality (Goal 5), Clean Water and Sanitation (Goal 6), Life Below Water (Goal 14), and Life on Land (Goal 15), are primarily addressed through broader socio-economic initiatives. In **6G-LEADER,** the focus is on enabling technologies that can indirectly support these ambitions by providing the connectivity foundations upon which such societal programmes can build.

Affordable and Clean Energy (Goal 7) is indirectly supported through innovations that significantly improve network energy efficiency and promote architectures that can integrate better with renewable energy systems. Similarly, Decent Work and Economic Growth (Goal 8) and Industry, Innovation and Infrastructure (Goal 9) are addressed through the deployment of new wireless infrastructures and the creation of business models that foster technological growth, digital inclusion, and economic competitiveness. The project also contributes to Reduced Inequalities (Goal 10) by enabling wider access to digital services through affordable 6G technologies, and to Sustainable Cities and Communities (Goal 11) by promoting energy-efficient, low-latency wireless communications fundamental to smart city infrastructures. Furthermore, Responsible Consumption and Production (Goal 12) and Climate Action (Goal 13) are supported by the emphasis on energy saving, resource-efficient network deployments, and the reduction of emissions through more intelligent and adaptable communication networks.

Although **6G-LEADER** does not directly address Life Below Water and Life on Land, it indirectly influences Peace, Justice and Strong Institutions (Goal 16) by promoting open, secure, and trustworthy network infrastructures that enhance institutional resilience. Finally, Partnerships for the Goals (Goal 17) are actively pursued through the project's strong commitment to collaboration with standardisation bodies, industry leaders, and academic institutions, ensuring that the technological advances are not developed in isolation but contribute to the broader 6G ecosystem. Table 3.1, maps the SDG towards the project contributions.

Table 3.1. Sustainable Development Goals mapped into 6G-LEADER.

#	Goal Name	Goal Description	Contribution of 6G-LEADER
1	No Poverty	End poverty in all its forms everywhere	N/A
2	Zero Hunger	End hunger, achieve food security and improved nutrition, and promote sustainable agriculture	N/A
3	Good Health and Well-being	Ensure healthy lives and promote well-being for all at all ages	Supports healthcare verticals via enhanced connectivity and low-latency applications for remote surgery and health monitoring
4	Quality Education	Ensure inclusive and equitable quality education and promote lifelong learning opportunities for all	Enables immersive education experiences through digital twinning and augmented/virtual reality
5	Gender Equality	Achieve gender equality and empower all women and girls	N/A
6	Clean Water and Sanitation	Ensure availability and sustainable management of water and sanitation for all	N/A
7	Affordable and Clean Energy	Ensure access to affordable, reliable, sustainable and modern energy for all	Indirectly supports through energy-efficient 6G technologies and promoting renewable energy integration

8	Decent Work and Economic Growth	Promote sustained, inclusive and sustainable economic growth, full and productive employment and decent work for all	Promotes new business models, digital inclusion, and economic growth through innovative wireless infrastructure
9	Industry, Innovation and Infrastructure	Build resilient infrastructure, promote inclusive and sustainable industrialization and foster innovation	Directly supports by advancing 6G wireless technologies and fostering standardisation
10	Reduced Inequalities	Reduce inequality within and among countries	Supports digital inclusion and wider accessibility through affordable and widespread 6G connectivity
11	Sustainable Cities and Communities	Make cities and human settlements inclusive, safe, resilient and sustainable	Enables smart city developments with low-latency, energy-efficient wireless networks
12	Responsible Consumption and Production	Ensure sustainable consumption and production patterns	Promotes energy efficiency and material recycling in network equipment
13	Climate Action	Take urgent action to combat climate change and its impacts	Supports climate change mitigation through energy-efficient network operations and reduced carbon footprint
14	Life Below Water	Conserve and sustainably use the oceans, seas and marine resources for sustainable development	N/A
15	Life on Land	Protect, restore and promote sustainable use of terrestrial ecosystems, sustainably manage forests, combat desertification, and halt and reverse land degradation and halt biodiversity loss	N/A
16	Peace, Justice and Strong Institutions	Promote peaceful and inclusive societies for sustainable development, provide access to justice for all and build effective, accountable and inclusive institutions at all levels	Indirectly supports through promoting open and secure networks
17	Partnerships for the Goals	Strengthen the means of implementation and revitalize the global partnership for sustainable development	Strong focus on collaboration with standardisation bodies, industry, academia, and public sector.

Table 3.2 outlines the Innovation Pillars of the **6G-LEADER** project, which represent the building blocks that organise the project's research, development, and validation activities. These pillars

capture the project's vision for the future of 6G networks, addressing critical challenges across the physical layer, radio access management, spectrum utilisation, and architectural intelligence. Each pillar reflects a forward-looking interpretation of what 6G systems will require—namely, dynamic reconfigurability, AI-native control loops, energy and spectrum efficiency, and a deep integration between semantic context and data transmission. By organising innovation around these key areas, **6G-LEADER** builds a clear path that connects its technology goals with real-world impact, while also addressing the technical needs expected in future 6G networks.

These pillars also serve as a bridge between the project's overall vision and its specific technical goals. For example, the focus on real-time control in O-RAN, or on coexistence across different frequency bands, highlights the project's effort to build practical, future-ready solutions. The pillars help to define the **6G-LEADER** baseline and lead to the definition of the main objectives, which describe what the project aims to achieve within each of these technical areas.

The goals of **6G-LEADER** are built around the challenges and ambitions defined by its innovation pillars. These objectives provide a focused direction for the project, outlining what needs to be developed, demonstrated, or improved to move 6G technologies forward. As described in Table 3.3, one objective focuses on building more intelligent and energy-aware signal processing at the physical layer using AI/ML techniques. Others tackle challenges like making spectrum access more efficient or designing semantic-aware communication systems that avoid transmitting redundant or unnecessary data. Some objectives are more architectural and operational in nature, such as enabling more flexible and automated RAN control using xApps, rApps, and the so-called dApps, which are applications intended to enable faster response times than xApps and rApps already in the O-RAN framework. There's also a strong focus on validation: several objectives are specifically aimed at demonstrating real-world impact through testbed deployments and PoC. This ensures that the project delivers practical, working solutions that can be tested, measured, and refined.

Table 3.2. Innovation pillars of 6G-LEADER.

Id	Innovation Pillar	Description
IP.1	AI/ML-driven Physical Layer	Enhancing prediction and optimization of network parameters.
IP.2	AI/ML-powered Multiple Access and Wireless for AI	Improving spectral efficiency through predictive access schemes.
IP.3	Highly Reconfigurable RF Components	Leveraging fluid antennas (FAs) and reconfigurable intelligent surfaces (RISs).
IP.4	Optimum Spectrum Usage and FR1-FR3 Coexistence	Addressing spectrum scarcity while maintaining coverage and efficiency.
IP.5	Semantics-Empowered Communications	Introducing AI-driven, goal-oriented data transmission and resource management.
IP.6	Real-Time RAN Control Loop	Embedding intelligence between O-RAN components for faster decision-making.
IP.7	Conflict Manager for xApps/rApps	Ensuring stability and coordination in O-RAN environments.

These objectives act as a link between the innovation areas and the expected outcomes of the project. They also form the basis for defining the technical KPIs, which help measure how well each objective is being achieved. To evaluate whether each objective is being achieved, the project defines a set of measurable KPIs that reflect the expected impact in technical terms. These KPIs are tightly linked to the objectives: for every goal set by the project, there are one or more KPIs that translate it into specific, quantifiable targets. For example, improving the physical layer using AI/ML models is supported by KPIs that aim to reduce communication overhead by 30%, cut latency in half, or improve spectral efficiency by 50%. These are not abstract targets—they are chosen to reflect realistic performance improvements that can be demonstrated and verified. These KPIs has a baseline for the traditional 5G implementations. By doing so, these KPIs can be evaluated with actual implementation or known data sources. Each baseline discussion will be discussed in detail in WP7.

Table 3.3. Objectives of 6G-LEADER.

Id	Objective	Description
O.1	AI/ML-empowered, predictive, and resource-efficient 6G PHY evolution	-Develop AI/ML-based models for real-time channel prediction and optimization. -Design AI/ML-driven transmitter and receiver chains to enhance performance. -Reduce communication overhead, energy use, and latency.
O.2	AI/ML-driven multiple access & Wireless for AI	-Develop random and non-orthogonal multiple access schemes. -Implement grant-free access schemes for enhanced spectrum efficiency. -Design over-the-air computing (AirComp) schemes for federated learning at the edge.
O.3	Highly reconfigurable RF components & FR1-FR3 coexistence	-Enhance fluid antenna (FA) and reconfigurable intelligent surface (RIS) capabilities. -Optimize spectrum usage with adaptive antenna reconfiguration. -Improve energy efficiency and spectral efficiency through novel RF designs.
O.4	Semantics-empowered 6G communications	-Implement goal-oriented and information-centric networking. -Reduce redundant packet transmissions and control overhead. -Develop AI/ML models to enable semantic-aware RAN resource allocation.
O.5	ML-driven O-RAN with xApps/rApps/dApps	-Deploy xApps, rApps, and dApps for automated RAN control. -Implement a Conflict Manager for O-RAN applications. -Reduce control overhead and latency in network operations.
O.6	Develop PoCs to validate the **6G-LEADER** RAN design	-Deploy five PoCs demonstrating project innovations. -Integrate **6G-LEADER** technology in large-scale testbeds. -Ensure compatibility with future SNS-JU projects.
O.7	Impact creation, standardization, and industry adoption	-Contribute to 3GPP, O-RAN Alliance, ITU-T, and ETSI standardization. -Develop open-source solutions and PoC documentation for SNS-JU integration. -Promote European leadership in 6G wireless technology.

As described in Table 3.4, the one-to-one or one-to-many mapping between objectives and KPIs ensures traceability and clarity throughout the project. It allows each area of innovation to be evaluated based on solid evidence, and it prepares the ground for validation in testbeds and real environments. Some KPIs focus on system-level performance, such as increased spectrum efficiency or reduced EMF exposure, while others target more functional aspects like the deployment of AI-enabled applications or the resolution of RAN control conflicts. These KPIs will later be used during the design and execution of the PoCs to check how the solutions perform under realistic conditions. In that way, they connect the high-level intentions of the project with hands-on validation activities, closing the loop between vision and implementation.

Table 3.4. Technical KPI mapped into the 6G-LEADER objectives.

Id	KPI Description	Objectives
KPI 1.1	Reduce communication overhead by 30% compared to SotA AI/ML algorithms	O1
KPI 1.2	Reduce end-to-end latency by 50% through efficient prediction mechanisms	O1
KPI 1.3	Improve spectral efficiency by 50% through distributed resource allocation	O1
KPI 2.1	Increase spectral efficiency by 40% using AI-driven multiple access schemes	O2
KPI 2.2	Reduce energy consumption by 30% in multiple access schemes	O2
KPI 2.3	Ensure spectrum requirements remain independent of the number of participating nodes in Wireless for AI	O2
KPI 3.1	Reduce EMF exposure by 30% using reconfigurable RF components	O3
KPI 3.2	Reduce energy consumption by 40% compared to non-reconfigurable solutions	O3
KPI 3.3	Improve spectral efficiency by 40% over benchmarks without reconfigurable RF	O3
KPI 4.1	Improve Age of Information (AoI) and derivative metrics by 20%	O4
KPI 4.2	Reduce cost of actuation by 10% and other timing and importance metrics	O4
KPI 4.3	Reduce non-effective packet transmissions and associated resource usage (including energy) by 90%	O4
KPI 5.1	Develop at least 10 xApps for energy-efficient RAN, low-EMF exposure, Wireless for AI, and multiple access optimization	O5
KPI 5.2	Ensure Conflict Manager resolves more that 50% conflicts with minimal RAN performance degradation	O5
KPI 5.3	Reduce O-RAN control plane overheads by 30% through semantic-awareness	O5
KPI 5.4	Enable real-time dApps-based control loop operating in sub-10ms	O5
KPI 6.1	Deploy 5 large-scale PoCs across testbeds	O6
KPI 6.2	Validate **6G-LEADER** innovations through real-world deployments	O6
KPI 7.1	Contribute 20+ standardization efforts in 3GPP, ETSI, ITU, O-RAN	O7
KPI 7.2	Publish high-impact research papers and patents	O7

The KVIs identified in Table 3.5 focus on the societal and environmental value of **6G-LEADER** beyond its technical KPIs. For example, the healthcare KVI (KVI.1) translates the technical achievements of low-latency and reliable connectivity (KPIs like reduced E2E latency or increased reliability) into measurable improvements in public health outcomes. Similarly, KVI.2 on education shows how XR and digital twinning create new modes of inclusive learning, directly tied to SDG 4. Energy-related KVIs (KVI.3, KVI.5) align with SDGs on energy efficiency and climate action, reflecting how reconfigurable RF (RIS/FAs) and AI-driven RAN optimisation contribute to sustainability. Meanwhile, KVI.4 on digital inclusion reflects the project's social impact by ensuring accessibility, affordable connectivity, and reduced inequality. On the governance and collaboration side, KVI.6 highlights trust and security in open and standardised infrastructures, while KVI.7 captures the value of strong partnerships and knowledge-sharing across academia, industry, and standardisation bodies. Altogether, these KVIs provide a direct link between SDGs, project objectives, and the envisioned PoCs, ensuring that technical innovations are justified by measurable socio-economic and environmental impacts.

Finally, the structured mapping from SDGs, innovation pillars, project objectives, and technical KPIs creates a coherent framework that guides the design and implementation of the **6G-LEADER** PoCs, as described in Table 3.6. Each step in this top-down methodology—starting from societal and environmental challenges (Table 3.1), to identifying key innovation domains (Table 3.2), followed by concrete project objectives (Table 3.3), and quantifiable KPIs (Table 3.4) builds the foundation for testing and validating the technical advances proposed by the project.

Table 3.5. Major KVI associated to the 6G-LEADER project.

id	KVI description	SDG	Objectives
KVI.1	Improved access to remote healthcare via low-latency, reliable communication enabling telemedicine, remote surgery, and monitoring	SDG 3	O2, O4
KVI.2	Increased inclusiveness in education through immersive XR/VR-based learning experiences	SDG 4	O4, O6
KVI.3	Reduction of RAN operational energy footprint by deploying energy-efficient technologies and RF components	SDG 7, SDG 13	O1, O3
KVI.4	Promotion of digital inclusion and reduced inequalities via affordable and widespread 6G connectivity	SDG 10	O5, O7
KVI.5	Enabling smart, resilient, and low-carbon urban infrastructures through energy-efficient wireless networks	SDG 11, SDG 12	O3, O6, O7
KVI.6	Strengthened trust in digital infrastructure via secure, open, and standardised network design	SDG 16	O5, O7
KVI.7	Enhanced collaboration and cross-industry innovation through joint testbeds, standardisation, and open-source contributions	SDG 17	O6, O7

The PoCs serve as the practical realisation of this framework. They are designed to cover specific combinations of innovation pillars and objectives, and each is associated with a unique subset of KPIs that reflects the performance, energy efficiency, reliability, and intelligence envisioned in the **6G-LEADER** architecture. This mapping ensures that the proposed innovations are not only theoretically aligned with the goals and targets of the project but also measurable and verifiable in real-world experimental conditions. The following table presents the five PoCs, detailing their focus areas, testbeds, and their direct alignment with the project's technical pillars, objectives, and KPI targets.

Table 3.6. PoC evaluated on the KPIs of 6G-LEADER.

PoC#	Title	Focus	Testbeds & Platforms	Innovation Pillars	Objectives	KPIs
PoC#1	XR and UAV seamless real-time interaction	Real-time interaction between XR and UAVs; AI/ML-driven predictive communication; Semantics-aware video processing	CNIT Federated Testbed (ARNO in Pisa & S2N in Genoa), ACC dRAX Platform, XR, UAVs, Edge Nodes (Nvidia Orin, Bluefield), SRS CU/DU Platform	1, 2, 5	O2, O4, O6, O7	2.1, 4.2, 4.3, 5.3.
PoC#2	Enhanced Mobile Broadband Experience	6G O-RAN RU and DU for FR3 band; Energy- and cost-efficient hybrid beamforming; RIS-assisted beamforming	DICAT, MB RU Platform, SRS CU/DU Platform	1, 3, 4, 6	O3, O6, O7	3.1, 3.2, 3.3, 6.2.
PoC#3	AI/ML Trainable 6G RIC Conflict Manager	Energy-efficient O-RAN management; Real-time conflict mitigation in Near-RT RIC; AI/ML-based optimization for x/rApps	DICAT, ACC dRAX RIC, SRS DU Platform	6, 7	O1, O2, O5, O6, O7	1.1, 1.2, 5.2, 5.3, 5.4.
PoC#4	Wireless for AI based on AirComp and empowered by	Wireless AI computing with AirComp; Semantics-aware task offloading;	CNIT Federated Testbed (ARNO & S2N), Development of dApps, semantic communication,	2, 5, 6	O2, O4, O5, O6, O7	2.1, 2.3, 4.1, 4.2, 5.4.

	semantically aware d/xApps	Efficient in-network AI computation	AirComp algorithms, SRS CU/DU Platform			
PoC#5	AI/ML-aided enhanced multiple access integrating low-EMF FAs in the FR1/FR3 bands	AI-driven multiple access schemes; Fluid Antennas (FA) for energy efficiency; Coexistence of FR1 and FR3 bands	UC3M Testbed with FA prototype, spectrum/network analysers, liquid conductor materials	1, 2, 3, 4	O1, O2, O3, O6, O7	1.3, 2.1, 2.2, 3.1, 3.3, 6.1.

3.2 High level Use Cases

The **6G-LEADER** project defines a set of high-level use cases that guide its architectural innovations, system requirements, and technology validation strategies. These use cases capture future service needs across verticals, integrating AI-native communications, energy efficiency, and ultra-reliable low-latency performance. Each use case serves as a practical scenario where key enablers—like AirComp, semantic processing, or conflict-aware traffic steering—can be tested under realistic conditions to demonstrate their relevance for the evolution of 6G networks. For clarity and modularity, related use cases are organized into Use Case Groups (UCGs), each representing a specific technological focus.

3.2.1 UCG1: AirComp-Enabled Use Cases

AirComp, short for Over-the-Air Computation, is emerging as a key enabler in several 6G-relevant scenarios due to its unique capability to merge communication and computation at the physical layer [35]. Rather than transmitting individual signals for decoding, AirComp allows devices to transmit simultaneously in a way that directly computes a function of their data in the air, dramatically reducing communication latency, energy consumption, and signalling overhead. This paradigm is especially useful in use cases requiring massive connectivity, low-latency collaboration, or real-time aggregation of distributed data. The following subsections present three representative AirComp-based scenarios that are actively explored within the **6G-LEADER** context: distributed learning for federated AI, ultra-efficient wireless Internet of Things (IoT) sensing and control systems, and scalable, low-latency localisation for the IoT. These use cases showcase the versatility of AirComp as a foundation for future wireless intelligence and control systems, highlighting its integration potential with emerging AI-driven architectures.

3.2.1.1 Distributed Learning

Traditional Machine Learning (ML) algorithms rely on performing training and inference centrally, supported by more powerful cloud computing resources. More specifically, ML models are trained at centralized servers by collecting the datasets from dispersed devices, a process that has several drawbacks. For instance, conducting data transmission and storing it in the server consumes vast amounts of bandwidth and energy and increases latency. Besides, data generated at local devices might be privacy sensitive, which is violated when they are gathered at central locations. To overcome these challenges, distributed ML techniques have been developed to enable edge devices to train their models locally and collaboratively train a shared global model [36]. A well-known technique in this area is Federated Learning (FL), where edge devices send their local model updates over a multiple access channel to a server, applying an aggregation function to update the global model. Then, the server sends back to the edge devices the updated global model for updating, using local data [37]. FL ensures data gravity, as only model parameters are exchanged. It also reduces communication overheads and power consumption, since raw data are not transmitted, compared to centralized ML, which relies on transmitting a high volume of training data [38]. However, despite these advantages, limited radio resources create bottlenecks in implementing FL in wireless communication systems. This is primarily due to the high-dimensional model or gradient parameters that must be periodically exchanged between the server and many edge devices. Conventional orthogonal multiple access schemes are inefficient in this context, as they allocate orthogonal channels to each device, leading to increased training latency and reduced spectral efficiency. When the goal of distributed learning is to compute a weighted sum of local model updates, consistent with nomographic function computation, AirComp can be leveraged for efficient model aggregation [35].

In this approach, the edge server directly receives an aggregated version of the analog-modulated local models or gradients, which are simultaneously transmitted by the edge devices. This simultaneous transmission enables AirComp to reduce communication and computation latency by a factor proportional to the number of devices scheduled to transmit concurrently. Moreover, the use of non-orthogonal communication allows multiple devices to share the same resource block, thereby increasing spectral efficiency [39]. In terms of energy expenditure, AirComp minimizes the idle time that edge devices would otherwise spend waiting for their transmission slot in orthogonal systems, reducing energy waste and accelerating the completion of each distributed learning epoch [40]. Further energy gains can be achieved when AirComp is integrated with techniques such as device selection based on optimal channel conditions, which allows for better performance with lower transmission power, and with optimization algorithms under power constraints, ensuring efficient operation in resource-limited environments [41]. Another domain that benefits from AirComp is privacy. Even when raw data remains on edge devices, privacy concerns persist due to advanced model inversion attacks capable of inferring local training data from shared model updates. AirComp mitigates this risk by enabling only aggregated updates to be received at the server. Since local model parameters are superimposed with those from other simultaneously transmitting devices, an eavesdropper gains access only to a composite signal, making it significantly harder to isolate individual contributions. Furthermore, privacy can be enhanced through differential privacy techniques, which introduce controlled perturbations into the aggregated model to preserve key statistical properties while obscuring the impact of any individual data point. Notably, in AirComp, channel noise naturally acts as a source of random perturbation, effectively supporting differential privacy and reducing the need for adding artificial noise at the local device level [42].

3.2.1.2 IOT and Wireless Control Systems

The Internet of Things (IoT) enables novel services by providing ubiquitous connectivity for sensors and machines, marking a shift from human-type communication to machine-type communication. However, achieving massive and reliable connectivity presents significant challenges, particularly regarding the scalability of radio resources, which may overwhelm the capacity of existing communication infrastructures. To address this challenge, Wireless Data Aggregation (WDA) has emerged as a promising solution for applications involving massive numbers of edge nodes performing distributed data measurements and transmitting them to an edge server for further processing [39]. To enable efficient WDA, AirComp has been proposed as a technique that allows multiple devices to transmit simultaneously over shared resource blocks. Unlike Orthogonal Multiple Access (OMA) [39], which assigns separate channels to each device, AirComp provides a spectrally efficient method for aggregating the signals of multiple sensors. In this context, a key use case for AirComp is distributed sensing, where densely deployed sensors monitor the physical environment and collectively create a digital representation of it [35]. For instance, services in environmental monitoring or smart cities, based on temperature and humidity measurements, often aim to compute a global function (e.g., average or maximum) across all sensors rather than collecting individual raw data. Similarly, in disaster prevention scenarios, the function of interest may be the maximum detected temperature or chemical concentration. AirComp is particularly well-suited for distributed sensing, as it enables simultaneous transmissions from all sensors while directly computing the desired function over the air, significantly reducing communication overhead and latency.

Another important application of AirComp lies in wireless control systems, which are essential in domains such as smart industries and agriculture, where groups of agents must collaboratively perform tasks [42]. In such systems, each agent iteratively gathers information from others to update its own state and achieve consensus. This process involves both a communication phase, where agents exchange information, and a computation phase, where each agent updates its state based on a function, typically the average of the other agents' states. This operation is crucial in IoT applications such as vehicular platooning and swarm UAV or robot formation control. In vehicular platooning, for instance, all participating vehicles must reach a consensus on key driving parameters, such as velocity, acceleration, and trajectory, to ensure coordinated movement. Given the mission-critical nature of such applications, achieving ultra-low latency is essential. AirComp addresses this need by significantly reducing per-round communication latency and accelerating the convergence process, leveraging its ability to perform computation concurrently with communication among distributed agents.

3.2.1.3 Distributed Localization in IoRT Systems

In the Internet of Robotic Things (IoRT), autonomous robots must continuously sense their environment, localize targets, and transmit pre-processed sensing data to access points (APs)[43]. Traditional systems treat communication and computation as separate processes, leading to increased spectrum consumption and communication delays. AirComp offers an efficient alternative, especially for latency-sensitive tasks such as target localization.

In one scenario, each robot senses a target and locally pre-processes its data. Instead of transmitting discrete packets, robots simultaneously transmit their processed data over a shared analog wireless channel using AirComp. The superimposed signal benefits from averaging effects, which mitigate the impact of noisy measurements from individual robots. This leads to improved spectral efficiency, reduced latency, and higher localization accuracy, which are crucial for applications such as surveillance, security, and responsive services in connected environments [44].

From another perspective, AirComp supports localization based on Anchor Nodes (ANs) that are aware of their precise positions. A mobile Server Node (SN) aims to estimate its location using feedback from the ANs without accessing their raw data. Each AN evaluates whether a specific cell in the area grid is a likely candidate for the SN's location, based on its known coordinates and the estimated SN distance. These binary votes are then transmitted simultaneously using AirComp. The SN receives the aggregated signal and identifies the Majority-Vote (MV) cell, representing its estimated location. This approach allows highly scalable, privacy-preserving, and efficient localization [45].

3.2.2 UCG2: XR & UAV seamless interaction

Figure 3.1 shows a reference scenario including XR headset and two UAVs. A neuromorphic camera, guaranteeing low energy consumption at the cost of lower resolution, is used to detect relevant events. UAVs exploit high-fidelity cameras that provide greater transmission capabilities and, consequently, consume more power. To optimize efficiency, the UAVs with high-fidelity camera are activated only when the neuromorphic camera detects a relevant event. Semantics-aware, real-time video transmission is then carried out to the XR headset. Specifically, semantic segmentation is applied to isolate and process only critical visual information, minimizing the data volume transmitted.

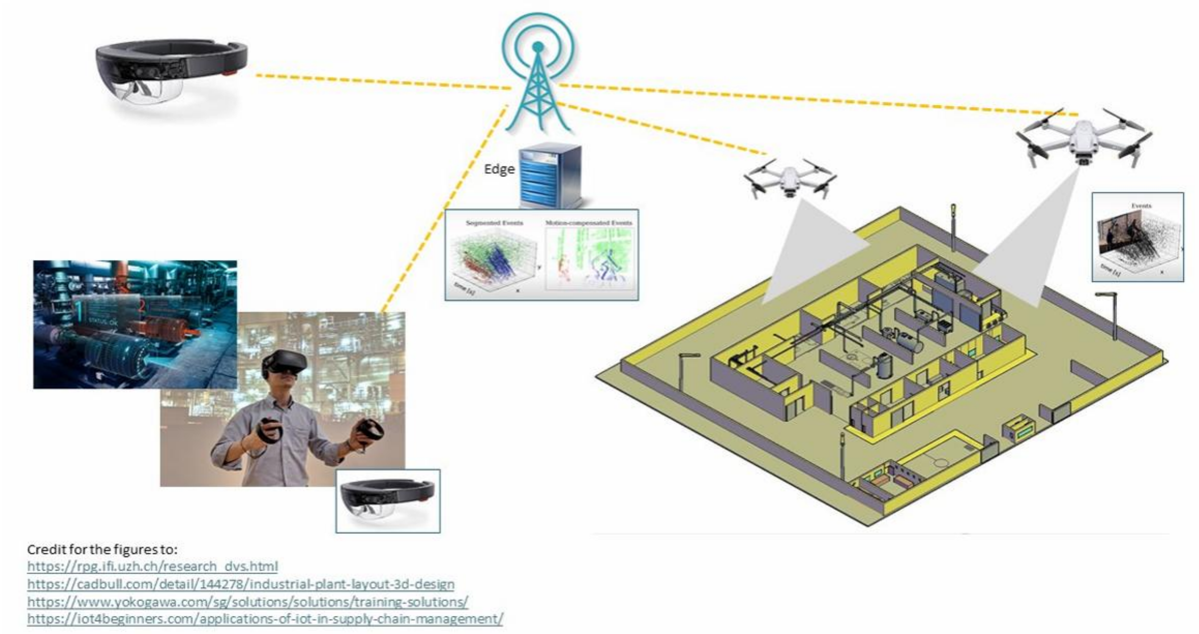

Figure 3.1. XR & UAV Seamless interaction.

A split inference approach is adopted, where part of the processing is done onboard the UAV, and the remaining is offloaded to an edge node via compressed feature vectors. This enables a responsive and power-efficient reconstruction of the surrounding environment, ensuring a high Quality of Experience (QoE) for the end user. Given the mobility of UAVs and the potential for ultra-low-latency requirements, the **6G-LEADER** framework is applied to enhance the communication link through advanced AI/ML-driven physical layer radio technologies, supporting high-throughput, bi-directional data exchanges during peak loads.

Accurate localization is also critical in this setup. Alongside video streams, the UAV's position must be transmitted to align spatial content with the user's visual perspective in the XR headset. Therefore, precise and timely localization is essential to reduce transmission overhead. By relying on preloaded environmental data and static maps within the user device, redundant data transfers are minimized, allowing smooth and immersive real-time interaction without overwhelming the network with unnecessary information.

3.2.2.1 XR-Assisted Infrastructure Inspection with UAVs

In this use case, UAVs equipped with complementary sensors support XR-assisted inspection of bridges, pipelines, or power lines. A neuromorphic camera continuously scans the structure at low energy cost, while a high-resolution camera is activated only when anomalies are detected. The UAV transmits semantics-aware, filtered visual data to an XR headset used by the operator, allowing real-time immersive inspection without overwhelming the network with redundant video streams.

The system leverages split inference: part of the image analysis is performed on the UAV to detect relevant patterns, while heavier tasks are offloaded to nearby edge nodes. By combining local and edge processing, inspectors receive immediate visual feedback while maintaining high accuracy in anomaly detection. The approach ensures responsive and power-efficient operation, while enabling advanced maintenance workflows across large-scale infrastructure.

3.2.2.2 Immersive Emergency Response with UAV Support

This use case envisions first responders using XR headsets in combination with UAVs during emergency situations, such as search-and-rescue or disaster management. UAVs fly into hazardous areas and stream semantically segmented video to operators, highlighting only the most relevant features, such as human silhouettes or heat signatures. This reduces data traffic while preserving critical situational awareness in real time.

For precise coordination, UAVs transmit their position alongside processed visual features, ensuring the XR headset correctly aligns the spatial information. AI/ML-driven radio techniques optimise low-latency communication during peak loads, guaranteeing smooth interaction even when multiple UAVs are deployed. By combining immersive XR feedback, semantic communication, and UAV mobility, the system empowers rescue teams to make faster and better-informed decisions in challenging environments.

3.2.3 UCG3: Energy Efficiency and Traffic Steering Use Cases

Improving the Energy Efficiency of future RANs is now deemed to be crucial as it accounts for around 70% of the total energy consumption of the 5G networks [46], [47]. Within the RAN, the Radio Unit (RU) stands out as the primary energy consumer, representing 40% of the total energy usage in RAN—an amount greater than that of all other RAN components, including air conditioning systems [48]. This needs to be enhanced to assist the network OPEX and improve environmental sustainability. Also, several 6G services will require advanced Traffic Steering capabilities to support various services such as mission-critical and prioritised services, e.g. for Public Protection and Disaster Relief (PPDR) first responders and public safety. Hence, for the 6G O-RAN networks, it is important to explore the gains that could be offered by the Traffic Steering xApps as their practical utilisation within the networks allows service providers to improve the user experience for such services. For O-RAN networks, WG1 specifications provide the description of use cases that are made possible or enhanced as a result of introduction of RIC within O-RAN architecture, including the energy efficiency and traffic steering use cases [49], [50].

The scope of **6G-LEADER** PoC activities include the evaluation, validation, and demonstration of two use cases for Energy Efficiency (switching on/off the cells) and Traffic Steering (to manage the traffic and achieve balanced cell load), which are of particular interest for the future 6G networks. Initially, the two Energy Efficiency and Traffic Steering use cases will be implemented and run individually on the advanced 5G network with multiple cells and users to allow evaluation and validation of practical gains they can offer to the network based on the designed xApps/rApps. Then, they will be run concurrently, which creates conflicting objectives for running AI/ML-driven optimisations within xApps/rApps of individual use cases. This provides a more realistic scenario for the future 6G O-RAN networks that are deemed to utilise multiple xApps/rApps with conflicting objectives for different use cases. For the Energy Efficiency and Traffic Steering use cases, the Conflict Manager unit within the RIC will conduct coordination and optimisations needed for the system to operate in a way that the objectives for individual applications and related policies will be supported, while it will also try to minimise the discrepancy that comes out as a result of conflicting actions, e.g. saving energy across the cells with negligible impact on user experience.

The innovations developed in this PoC demo have potential impact on several verticals, mainly for the use cases that require adjustments in the network planning due to special events, such as emergency services (e.g., emergency response and disaster management) and during network fault management. Conflict management solutions from this PoC will also contribute to fully autonomous zero touch network evolution where conflict management solutions will enable conflicting xApp/rApps to run in the network autonomously.

3.2.3.1 Energy-Aware RAN Optimisation

In this use case, the RAN dynamically reduces its energy footprint by selectively switching off underutilised cells during periods of low traffic, while ensuring service continuity. AI/ML-driven xApps running in the Near-RT RIC predict traffic demand based on historical data and real-time monitoring, allowing proactive control of radio units. When load increases, sleeping cells are reactivated seamlessly, ensuring that users experience no service disruption.

The approach directly addresses network OPEX and environmental sustainability by reducing unnecessary energy consumption in the RU, which accounts for a major portion of RAN power usage. Through conflict-aware orchestration, the system balances energy-saving actions with QoE requirements, ensuring that cost reductions do not come at the expense of coverage or reliability. This makes it particularly relevant for dense urban deployments and large-scale operators seeking to minimise their carbon footprint.

3.2.3.2 Conflict-Aware Traffic Steering for Mission-Critical Services

This use case focuses on enhancing service quality in scenarios where different user groups have competing requirements. For instance, during an emergency event, public safety communications may need priority over regular broadband traffic. Traffic Steering xApps classify flows by priority and redistribute users across available cells to guarantee QoS for mission-critical services. Real-time adjustments are made within the Near-RT RIC, supported by a Conflict Manager that resolves inconsistencies when multiple xApps propose conflicting actions.

By combining AI-powered steering policies with semantic tagging of traffic intent, the system achieves balanced resource allocation even under stress conditions. It guarantees that first responders benefit from reliable and low-latency connections, while commercial users still maintain acceptable service levels. This conflict-aware approach provides a realistic path towards fully autonomous, zero-touch networks that can self-optimise for diverse requirements without manual intervention.

3.2.3.3 Adaptive RU Energy Management in Multi-Cell RANs

This use case addresses the challenge of reducing energy consumption at the RU level, which accounts for nearly 40% of total RAN power usage. The system monitors traffic conditions across multiple cells and dynamically adjusts RU operation modes, such as low-power transmission, antenna muting, or partial deactivation of transceiver chains. When network demand rises, RUs are quickly reconfigured back to full capacity, minimising service degradation.

AI/ML models running in the Near-RT RIC continuously evaluate user distribution, mobility patterns, and QoE metrics to identify opportunities for energy savings. By embedding decision-making in the RIC, the approach ensures sub-10ms responsiveness to traffic fluctuations, enabling fast transitions between active and energy-saving states. This mechanism directly reduces operational costs while improving sustainability, making it a cornerstone of future large-scale 6G deployments where energy efficiency is as critical as throughput and latency.

3.2.4 UCG4: Reconfigurable Surfaces and Antenna Use Cases

The reconfigurability of FA and RIS technologies enable a dynamic adaptation to changing channel conditions and thus maintain reliable connectivity.

3.2.4.1 RIS - increased coverage with FR3 hybrid beamforming

Since the channel characteristics depend heavily on the carrier frequency, it focus in the following on the FR3 band, particularly the 7-8GHz range. To achieve the same coverage as in FR1 systems while reusing existing cell sites, the number of antenna elements at the base station must increase by at least a factor of four in FR3. This makes full digital beamforming expensive and energy-consuming, as every antenna element requires a dedicated transceiver chain. Therefore, hybrid beamforming antenna systems must be considered for base stations with many antennas in order to maintain the same antenna aperture as FR1 systems and achieve the required higher number of antenna elements. One promising candidate technology for hybrid beamforming in this regime are Near-Field Fed Reflective Intelligent Surface (NFED-RIS) systems. In particular, a small number of active antennas and a large number of passive antennas at the RIS illuminated by the active antennas in the near field are deployed to achieve a large antenna aperture. Each active antenna is equipped with a transceiver chain and the passive antenna elements of the RIS are controlled by phase shifters. Notice that the number of passive antennas determines how narrow the beam can be, and the number of active antennas defines the number of beam directions. Therefore, a large number of passive antennas results in highly directive beams. The considered system in PoC #2 with four active antennas and an 8x8 RIS array can transmit one beam. Multiple beams to one or several users can be supported by stacking several NFED-RIS systems, resulting in a hybrid BF system. Compared to a fully digital beamforming system in FR1, a larger beamforming gain due to the large number of passive antennas is obtained, leading to higher signal-to-noise ratios and spectral efficiency. Moreover, the simple RIS hardware technology leads to higher efficiency in terms of cost and energy in comparison with fully digital beamforming [51].

3.2.4.2 Fluid antennas and blockage mitigation

Energy efficiency and EMF exposure have been denoted as major concerns addressed by **6G-LEADER**. Accordingly, FA use case aim to take advantage of its reconfigurability to mitigate or avoid potential fluctuations of the communication channel. In particular, FA can be installed on the UE to help reducing the effects of fading and/or blocking events. If an obstruction like a hand or a body would impair the link's performance, an FA might dynamically change its physical position to avoid being impacted. For FR1 or FR3 frequency range, the impact of these types of obstacles in communications is critical (>10 dB), and the distance (in terms of electrical wavelengths) is small allowing a fast re-positioning of the FA.

Since the energy required to move the antenna is often low, eliminating the aforementioned obstruction situations would obviously improve antenna efficiency, data rate, and overall energy usage. Moreover, electromagnetic radiation exposure from the antenna would be significantly decreased if the obstacles were human persons or bodily parts. Thus, reducing EMF-exposure in comparison to traditional antenna systems.

3.3 Use cases analysis

This section provides an analysis of the high-level use cases defined in the **6G-LEADER** project, structured around thematic UCGs. The analysis emphasizes how each use case utilizes targeted innovations across the **6G-LEADER** architecture to fulfil functional requirements and advance the project’s overall performance objectives.

3.3.1 UCG1 Analysis

The UCG1 considers federated learning, IoT sensing and control, and distributed localization, demonstrating AirComp's versatility across domains requiring low-latency, energy-efficient, and privacy-preserving data aggregation.

From the PHY/MAC perspective, UCG1 emphasizes that AirComp exploits simultaneous analog transmissions together with pre-equalization and grant-free multiple access. These techniques help mitigate channel impairments and allow efficient aggregation to take place directly at the physical layer, while the MAC layer coordinates the sharing of resources.

The key architectural elements engaged are:

- *RT and Near-RT Control Loops:* AirComp-enabled transmissions require fine-grained coordination of simultaneous analog transmissions, which is managed through dApps at the O-DU/O-CU level. Scheduling, synchronization, and RF control occur within sub-10 ms latency budgets, especially in vehicular control or robot localization scenarios.
- *xApps in Near-RT RIC:* Oversee AI-driven functions such as device selection, adaptive aggregation policies, dynamic power control, and aggregation function adaptation. In distributed learning, for example, xApps can prioritize devices with favourable channels to enhance learning efficiency.
- *Semantic-Aware Modules:* Exploit metadata (e.g., task descriptors, relevance tags) to assign transmission resources based on data importance, e.g., critical control inputs vs. background sensing data. This enhances spectral efficiency and enables resource reuse across coexisting services.

Requirements addressed by *UCG1*:

- Distributed ML model updates via AirComp support federated intelligence at the edge.
- AirComp reduces idle time, minimizes scheduling overhead, and enables concurrent analog transmission.
- Aggregation over the air hides individual data contributions and supports differential privacy mechanisms.
- Supports latency-sensitive scenarios such as vehicular consensus and IoRT localization.
- Non-orthogonal AirComp allows multiple devices to operate simultaneously on shared spectrum.
- Enables task-aware prioritization of aggregated data based on application intent.

3.3.2 UCG2 Analysis

UCG2 investigates an immersive XR application powered by the collaborative operation of neuromorphic sensors and UAVs equipped with high-fidelity cameras. It aims to enhance visual perception, communication efficiency, and system responsiveness in fast-changing environments, such as infrastructure inspections, training simulations, or emergency response scenarios. The scenario combines advanced sensory technologies, semantic data reduction, and AI/ML-driven network optimization, positioning it as a representative example of real-time, mission-critical services envisioned in 6G.

From the PHY/MAC perspective, UCG2 focuses on enabling low-latency video streaming over dynamic links, where UAV mobility places stringent demands on beam management, synchronization, and the handling of short-packet transmissions. Semantic compression alleviates the physical-layer load, while MAC-layer scheduling combined with split inference ensures that critical features are delivered on time despite varying channel conditions and interference.

The Key architectural elements engaged are:

- *RT and Near-RT Loops:* dApps on the UAV/DU enable real-time sensing control, segmentation, and PHY-level coordination. xApps manage adaptive scheduling, beam alignment, and compression strategies based on user-centric KPIs and UAV position.
- *Semantic-aware processing:* Task descriptors prioritize transmission of critical features, leveraging semantic compression to reduce network load.
- *Split inference orchestration:* AI/ML orchestration modules coordinate task offloading and ensure edge inference responsiveness.
- *Non-orthogonal PHY tuning*: AI-enhanced multiple access and waveform reconfiguration optimize link reliability and spectral efficiency during mobile XR interaction.

Requirements addressed by UCG2:

- Split inference and adaptive PHY optimization are managed by AI-based control.
- Event-triggered UAV activation and semantic compression minimize unnecessary transmission.
- dApps support ultra-low latency streaming and alignment of sensory and positional data.
- UAV trajectories and user perspectives are synchronized using AI-enhanced localization and adaptive beamforming.

3.3.3 UCG3 Analysis

UCG3 responds to the increasing demand for intelligent conflict resolution within the RAN, especially in scenarios involving competing objectives like energy efficiency and service prioritization. In future 6G O-RAN ecosystems, various AI/ML-powered xApps and rApps will operate concurrently, each targeting different optimization goals. PoC#3 showcases how these goals can be effectively harmonized using a trainable Conflict Manager embedded in the Near-RT RIC, enabling dynamic, policy-driven adaptation of RAN behaviour.

From the PHY/MAC perspective, UCG3 shows that conflict resolution directly shapes radio performance and access control. Energy-aware RAN optimization, achieved through switching RU components, muting antenna chains, and adjusting transmission power, influences SINR, coverage, and reliability. At the MAC layer, traffic steering relies on dynamic scheduling, prioritization, and handover control to protect mission-critical services, ensuring that energy saving and prioritization strategies do not compromise link quality.

The Key architectural elements engaged are:

- *Near-RT RIC:* Hosts both the Energy Efficiency and Traffic Steering xApps and the Conflict Manager, which consists of:
    - A Conflict Detection Module to identify rule or policy conflicts.
    - A Resolution Engine to prioritize or recompose actions using AI/ML models trained on historical data.
    - A Conflict Database that retains resolution history, policy hierarchies, and optimization outcomes.
- *Non-RT RIC:* Trains the underlying AI/ML models and enforces long-term policy alignment across use cases.
- *Semantic tagging and task descriptors*: Used by xApps to classify and signal the intent and criticality of control decisions, aiding the Conflict Manager in reasoning about policy implications.
- *Near Real-time feedback loops*: xApps monitor RAN performance post-resolution to update decision policies.

Requirements addressed by UCG3:

- The Conflict Manager operates using learned resolution strategies, enabling dynamic policy mediation.
- The Energy Efficiency xApp directly targets reductions in RU power consumption without compromising core services.
- Conflicting decisions may be resolved in the near-RT loop, ensuring low-latency reaction to system conditions; but can also be resolves in the Non-RT RIC via A1 policies implementation.
- Supports the evolution toward fully autonomous, zero-touch networks by embedding conflict resolution logic into the RAN control plane.

3.3.4 UCG4 Analysis

The UCG4 examines how RIS and FAs can enhance beamforming performance, improve spectral efficiency, and reduce energy consumption in mid-band 6G deployments. These technologies play a key role in meeting the eMBB targets of 6G, particularly in the 7–8 GHz FR3 band, by enabling beamforming architectures that are cost-effective, energy-efficient, and highly adaptable.

Leveraging RIS-based near-field fed architectures in the FR3 band enables the realization of large antenna apertures without the high cost and energy consumption associated with fully digital beamforming. By integrating a limited number of active transceivers with large passive arrays managed through phase shifters, the system achieves narrow, highly directive beams and substantial beamforming gain. This configuration improves SNR and spectral efficiency while keeping hardware complexity low. The PoC#2 setup showcases the ability to generate precise beams. Multi-user and multi-beam capabilities can be supported by stacking multiple such systems. Altogether, this hybrid beamforming approach offers a scalable, cost-effective solution for expanding FR3 deployments without increasing site density.

FAs introduce a physical reconfigurability mechanism that allows dynamic mitigation of obstructions, such as body blockage, which can cause signal attenuation exceeding 10 dB in FR1 and FR3 bands. When integrated at the UE, FAs can actively reposition to sustain optimal radiation paths, thereby enhancing link reliability and improving overall throughput. This adaptability minimizes unnecessary energy consumption and also improves antenna efficiency. Additionally, in environments where human presence obstructs the line-of-sight, the ability of FAs to reorient reduces EMF exposure, aligning with **6G-LEADER**'s goal of lowering EMF exposure relative to conventional static antenna systems.

From the PHY/MAC perspective, UCG4 shows that RIS phase control and FA port reconfiguration directly shape the radio channel, affecting SINR, coverage, and link robustness. At the MAC layer, these reconfigurable elements enable dynamic beam management and user scheduling, which support efficient multiuser service while lowering power consumption.

The Key architectural elements engaged are:

- *Near-RT RIC:* Manages hybrid beamforming optimization through xApps that dynamically select RIS configurations and adjust beam directions in response to real-time channel conditions and user mobility patterns.
- *Non-RT RIC:* Trains long-term RIS control policies and FA repositioning strategies, leveraging historical network data and radio environment maps.
- *RT Control Loop / dApps:* Executes time-sensitive commands for RIS phase adjustments and FA reconfigurations, enabling rapid link adaptation in response to changing network conditions.
- *AI/ML Orchestration:* Coordinates interactions between active and passive antenna elements and determines optimal beam configurations based on SNR, user density, and coverage objectives.
- *Reconfigurable Radio Hardware:* Integrates RIS modules at the BS and FAs at the UE, enabling low-power and adaptable antenna behaviour.

Requirements addressed by UCG4:

- Spectral efficiency and beamforming gain are enhanced by the large passive aperture of the RIS and the precise optimization of phase control, enabling narrow, high-gain beams.
- Lower power consumption is realized by minimizing the number of active transceivers in RIS and by avoiding retransmissions through FA repositioning.
- The RIS hardware design reduces implementation complexity and cost compared to conventional fully digital beamforming systems.

- FA repositioning avoids radiating towards human obstructions, aligning with exposure mitigation goals.
- Real-time link adaptation is enabled through sub-10 ms control loops managing RIS phase shifts and FA repositioning in response to changing conditions.

3.4 Low level mapping

This section presents a detailed alignment between the KPIs outlined earlier in Table 3.4 and the relevant UCGs emerging from the project's PoCs. For each UCG, it highlights how specific architectural advancements, control strategies, and AI-enabled mechanisms collectively contribute to achieving the intended KPIs. The following tables (Table 3.7, Table 3.8, Table 3.9 & Table 3.10) describe the mapping between the KPI and the UCG1, UCG2, UCG3 and UCG4 respectively.

Table 3.7. Low level mapping for UCG1.

KPI ID	How It Is Addressed in UCG1 (PoC#4)
KPI 2.1	AirComp enables simultaneous transmissions from multiple devices using non-orthogonal access, enhancing spectral efficiency without increasing bandwidth.
KPI 2.3	The use of AirComp for over-the-air model aggregation decouples spectrum use from the number of nodes, supporting scalable wireless AI.
KPI 4.1	By computing results in-air and reducing communication delays, AirComp improves freshness of information crucial for time-sensitive decisions
KPI 4.2	Semantic-aware data aggregation and task relevance filtering reduce unnecessary actuation and control signalling.
KPI 5.4	Real-time control of sensing and aggregation is achieved through sub-10ms dApp loops managing analog transmissions and function computation.

Table 3.8. Low level mapping for UCG2.

KPI ID	How It Is Addressed in UCG2 (PoC#1)
KPI 2.1	AI-enhanced PHY-level tuning and non-orthogonal waveform design optimize XR-UAV link efficiency.
KPI 4.2	Split inference reduces transmission cost and task scheduling aligns with semantic relevance.
KPI 4.3	Semantic segmentation transmits only critical visual features, minimizing data overhead.
KPI 5.3	Semantic-aware dApps reduce signalling and optimize control decisions based on task intent.

Table 3.9. Low level mapping for UCG3.

KPI ID	How It Is Addressed in UCG3 (PoC#3)
KPI 1.1	Conflict Manager filters and prioritizes control actions, reducing redundant signalling between xApps.
KPI 1.2	Real-time conflict resolution avoids delays caused by xApp contention and ensures timely execution of control commands.
KPI 5.2	Conflict Manager enforces AI/ML-based resolution logic to maintain RAN KPIs while resolving conflicting policies.
KPI 5.3	Semantic tagging of xApp intent allows the Conflict Manager to streamline coordination and minimize unnecessary control messaging.
KPI 5.4	The Conflict Manager operates within the Near-RT loop, ensuring decisions meet strict real-time deadlines.

Table 3.10. Low level mapping for UCG4.

KPI ID	How It Is Addressed in UCG4 (PoC#2)
KPI 3.1	FAs installed on UEs dynamically adjust their orientation to reduce direct human exposure, thereby lowering EMF levels during operation
KPI 3.2	NFED-RIS architecture reduces the number of active transceiver chains needed for beamforming, minimizing energy consumption at the BS
KPI 3.3	Highly directive RIS-enabled beams and FA-based blockage mitigation maintain high-quality links, enabling efficient spectrum reuse and higher throughput
KPI 6.2	PoC#2 demonstrates real-world RIS and FA deployments in the FR3 band, validating reconfigurable hardware technologies under operational conditions

In summary, the low-level mapping across UCGs, PoCs, and KPIs establishes a clear connection between the project's experimental activities and its performance objectives. It ensures that the technical innovations developed in **6G-LEADER** remain aligned with measurable outcomes such as spectral efficiency, latency reduction, energy saving, and semantic-aware communication gains. Moreover, the mapping provides a transparent framework for tracking progress throughout the project while enabling future deliverables to incorporate experimental results and progressively refine the KPI-PoC alignment as the proposed solutions evolve and mature.

4 Advanced and AI-driven communication techniques

The growing complexity and scale of modern wireless networks demand communication methods that go beyond conventional data transmission approaches. Advanced techniques such as semantic communication and AirComp are redefining how information is exchanged and processed. Specifically, semantic communication focuses on transmitting the intended meaning rather than raw data, reducing communication overhead and enabling more efficient exchange of information. Moreover, AirComp exploits the superposition property of wireless channels to perform data aggregation directly in the air, offering ultra-low latency and energy-efficient solutions. These communication paradigms can be further enhanced by AI-driven techniques, which leverage learning in order to optimize resource allocation, channel estimation and/or interference management. The effective design and development of these techniques can pave the way for more intelligent, context-aware networks that are capable of supporting the diverse and dynamic requirements of 6G networks. In what follows, a state-of-the-art analysis is provided together with **6G-LEADER** advances related to these technologies.

4.1 AI/ML-aided Physical-Layer Evolution

The 3GPP Technical Specification Group had its first group-wide 6G Workshop to discuss the vision and priorities for next generation radio technologies, system architectures, core networks and protocols [52], [53]. One of the key objectives, relevant to WP3, is the evolution of MIMO technology with an increased number of ports and antenna elements over FR3 bands as well as the native integration of Artificial Intelligence / Machine Learning (AI/ML) frameworks into the network for intelligent automation, optimization and improved efficiency.

Overall, the integration of AI/ML into wireless communication systems is transforming the way networks are designed and optimized. These technologies are expected to play an integral role in addressing the growing complexity and dynamic nature of future 6G networks. The following subsections outline advanced AI/ML-driven techniques at the physical layer and highlight the novelty of **6G-LEADER** in this area. In particular, they focus on intelligent channel estimation, predictive modelling and adaptive parameter optimization, all of which are critical for achieving high data rates, low latency and energy efficiency in next-generation communication systems.

4.1.1 ML-aided channel learning and prediction

Acquiring accurate Channel State Information (CSI) is a key challenge in modern wireless networks due to two main factors: the high dimensionality resulting from massive MIMO and wideband configurations, and the rapid temporal variations in high mobility scenarios. Traditional methods such as Least Squares (LS) and Linear Minimum Mean Square Error (LMMSE) suffer from high pilot overhead and limited accuracy, making them unsuitable for emerging applications like high-precision localization. These challenges motivate the adoption of AI/ML techniques to reduce overhead, enhance estimation accuracy, and enable predictive CSI for dynamic environments. The project will adopt resource-efficient AIML methods with online adaptation and feedback mechanisms to sustain real-time operation and reduce signalling, directly supporting KPI 1.1 (−30% communication overhead) and KPI 1.2 (−50% E2E latency).

Beamforming is crucial in massive MIMO systems to transmit and receive energy and mitigate severe path loss. Traditional codebook-based methods rely on a predefined set of beamforming vectors and perform exhaustive searches over beam pairs. To reduce the overhead of beam training, recent studies propose using a multi-classifier network to predict the optimal beam pair based on a limited number of beam measurements [54]. To further improve beam alignment accuracy and eliminate quantization losses inherent in fixed codebooks, codebook-free beamforming approaches have been introduced, leveraging deep neural networks (DNNs) to directly generate the beamformer from received signals or coarse CSI [55], [56]. Channel charting will produce a low-dimensional latent map for robust beam management and user localization, while Bayesian optimization will steer beam selection with calibrated uncertainty; together, these improve spectral efficiency toward KPI 1.3 (+50%).

For wideband massive MIMO channel sparsity that exists in the delay-angle domain, make them well-suited for Compressed Sensing (CS)-based estimation. However, traditional CS algorithms, such as approximate message passing and sparse Bayesian learning, tend to converge slowly. To address this issue, deep unfolding techniques have been proposed, which transform each iteration of a CS algorithm into a neural network layer with learnable parameters [57], significantly accelerating convergence while preserving the underlying algorithmic structure. Alternatively, end-to-end deep learning architectures, such as autoencoders, provide a unified framework that can jointly optimize beamforming and channel estimation through data-driven learning [58], [59]. End-to-end and block-based AIML-empowered transmitter/receiver chains with online adaptation and feedback mechanisms will reduce communication overheads and improve robustness, supporting KPI 1.1 (−30%).

In terms CSI prediction, conventional methods based on AutoRegressive (AR) models often suffer from performance degradation due to model mismatch. Deep sequence models can address this limitation by learning temporal dependencies directly from data. Specifically, Recurrent Neural Networks (RNNs), such as Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) models, have been widely adopted for CSI prediction [60]. More recently, transformer-based attention networks have shown superior performance by capturing long-range dependencies and reducing error propagation through parallel processing of historical CSI [61]. In addition, generative models, such as Generative Adversarial Networks (GANs), Variational AutoEncoders (VAEs), and diffusion probabilistic models, are emerging as powerful tools, as they can learn the underlying distribution of CSI and synthesize future channel states [62], [63]. To handle non-stationary environments and distribution shifts, the project will combine channel charting with Bayesian methods (e.g., Gaussian Processes, Bayesian optimization) to obtain calibrated uncertainty for robust scheduling, power control, and beam selection; online adaptation of AIML models will mitigate performance degradation across environments, supporting KPI 1.2 (latency) and KPI 1.3 (spectral efficiency).

Applying AI/ML to channel estimation and prediction presents several key challenges. First, there is an inherent trade-off between accuracy and complexity: while advanced models can achieve high estimation performance, they often incur increased inference latency and computational demands, limiting their practicality in real-time systems. Second, robustness and generalization remain a persistent concern, as models trained on specific datasets often struggle to maintain performance when exposed to distribution shifts in new propagation environments. These challenges highlight the need for AI/ML models that are both efficient and resilient across diverse channel conditions. In addition, such capabilities directly support **6G-LEADER**'s broader innovation pillars: they enable AI/ML-aided multiple access and Wireless for AI via robust power/rate allocation and AirComp aggregation (KPI 2.1–2.3); inform RIS and fluid antenna configuration for spectrum and energy efficiency (KPI 3.2–3.3); and fit into real-time control loops with dApps for sub-10 ms network decisions (KPI 5.4).

4.1.2 ML-aided parameter optimization

Traditional model-based optimization techniques often struggle to scale with the increasing complexity of wireless systems, which are characterized by dynamic topologies, high mobility, and massive device connectivity. ML-aided parameter optimization can have an important role towards ensuring the efficient utilization of the network's resources with respect to scheduling, power control, spectrum allocation, and interference management. Moreover, ML-based techniques can be effective in the beamforming design through precoding optimization and in RIS-aided networks through the optimization of the phase-shifts. In what follows, some ML-aided approaches proposed in the literature are discussed.

4.1.2.1 Resource allocation and power control

Specifically, multi-agent Deep Reinforcement Learning (DRL) has been used for jointly optimizing dynamic channel access and power control, allowing autonomous decisions from the users on their transmission policy thus maximizing the sum-rate or achieving proportional fairness [64]. Furthermore, DRL-based algorithms for power control in multi-user cellular networks have been proposed [65], demonstrating their superiority over model-based methods in sum-rate performance as well as computational efficiency. A deep learning-based framework for resource allocation in Non-Orthogonal Multiple Access (NOMA) networks has also been investigated, focusing on user association, subchannel assignment and power allocation [66]. The proposed framework achieved high energy efficiency with low computational complexity. A deep learning framework has also been presented for optimizing resource allocation in multi-channel cellular systems with Device-to-Device (D2D) communication [67].The proposed framework maximizes the spectral efficiency of D2D pairs while guaranteeing a minimum rate for the cellular users.

In **6G-LEADER**, complementing DRL approaches, we advocate the incorporation of Gaussian Process Regression (GPR), which facilitates a probabilistic, non-parametric modelling framework for tasks such as interference prediction and resource forecasting. For instance, GPR has been successfully applied to predict interference in dense 6G networks, enabling proactive and uncertainty-aware resource allocation strategies [68]. Unlike black-box neural models, GPs naturally quantify predictive uncertainty, making them especially valuable for decision-making in scenarios with sparse or partially observed data.

4.1.2.2 Beamforming

DRL has also been applied for the hybrid beamforming design in full-duplex millimetre wave systems [69]. The optimization problem is modelled as a Markov Decision Process (MDP) towards maximizing the spectral efficiency while mitigating the self-interference. Moreover, a deep neural network-based hybrid beamforming scheme has been proposed for a massive MIMO system, formulated as an autoencoder neural network [70]. The scheme is based on self-supervised learning and outperforms conventional methods. DRL-based approaches have also been employed to optimize the RIS's phase shifts [71]. The control of the RIS is formulated as an MDP, and DRL is applied for real-time control of the phases, resulting in significant performance gains.

A hybrid model-based and data-driven framework for wireless systems has also been proposed [72]. The results demonstrate an improvement in convergence speed and the obtained solution is closer to the optimal one compared to the conventional model-free ML approach. More recently, GPR has been proposed as a predictive control tool for RIS-aided systems, where it models the response surface between phase configurations and system performance metrics. This allows for efficient exploration of the RIS parameter space and facilitates uncertainty-aware control strategies, particularly when integrated with reinforcement learning for data-efficient policy learning [73].

4.1.2.3 Approach in 6G-LEADER

To overcome the limitations of purely data-driven approaches, the project aims at developing hybrid frameworks that combine model-based optimization with ML techniques. For example, model-driven initialization followed by GP-based fine-tuning has shown improvements in convergence speed and robustness. Such approaches are especially promising for scenarios with limited training data or where interpretability and safety are essential, such as autonomous control loops in cyber-physical systems. Ongoing research focuses on integrating these techniques into unified, scalable frameworks capable of online learning and multi-task optimization under the highly dynamic conditions envisioned for 6G.

4.1.3 Data collection and lifecycle

Initial work on standardising AI/MLOps in the context of O-RAN has been done by O-RAN WG2. Typically, AI/MLOps comprises various main steps, of which the data collection and preparation step is the first one, being the others: AI model training, validation and publishing, deployment, AI/ML execution and inference, and continuous operations. Data is collected from the RAN infrastructure through A1/O1/E2 O-RAN interfaces, and, after processing and preparation, feed back to the O-RAN node for inference or to the AI/ML Model management for training. Data collected through the O1, A1 and E2 can be stored in large datasets to be extracted upon request. Indeed, this data can be used either at run time, e.g., for inference or to feed adaptive solutions, or offline, e.g., in the design, training and testing of AI/ML models. Different measurement data can be collected from the RAN over time, such as throughput, latency, or channel quality information.

The data preparation step is a preliminary step in which data is cleaned and formatted to fit the inference and training input format and requirements of the AI/ML mode that will be embedded in a x/r/dApp. Various operations can be performed, e.g. dimensionality reduction using autoencoders, as well as data processing procedures (normalisation, scaling and reshaping). These technologies can identify and correct errors, inconsistencies, and duplicates in large datasets.

Learning models at the RAN need to be kept updated to avoid the degradation of their performance. This is due mainly either to the dynamic evolving of data profiles describing the environment or to the very dynamic changes that may occur in the radio access environment. For this reason, data collection for training is not just one once a time task, but it is done continuously over time and is part of the MLOps automation pipeline.

In the context of 6G, the Network Data Analytics Function (NWDAF) - previously defined in the 5G architecture from 3GPP Rel. 15 - can perform the data collection from various Network Functions (NFs). This data is then used for various purposes, from network optimization to anomaly detection.

4.2 Multiple Access and Over-the-air Computation Schemes

Modern wireless networks face growing scalability challenges due to the increasing number of connected devices and the need for low-latency data processing. Traditional multiple access schemes rely on orthogonal resource allocation to mitigate interference, assigning separate time, frequency, or spatial resources to each device. While effective for moderate traffic, this approach becomes inefficient in dense scenarios such as IoT, edge intelligence, and distributed learning, where sequential transmissions introduce significant latency and energy overhead.

AirComp offers a paradigm shift by leveraging the superposition property of wireless channels to perform functional computations directly during transmission. By allowing concurrent transmission, AirComp enables the aggregation of distributed data within the physical layer, eliminating the need for separate data collection and reducing both latency and energy consumption. This integrated communication-computation approach provides a scalable solution well-suited for emerging AI-native applications and federated learning frameworks operating at the network edge.

In the following subsections, the concepts of AirComp and both random and non-orthogonal multiple access schemes, as envisioned within the **6G-LEADER** framework are introduced.

4.2.1 Over-the-air computation (AirComp) schemes

As 6G architectures evolve toward data-centric designs, the focus is shifting from merely transmitting raw, unprocessed data to performing computations during transmission. This paradigm is especially beneficial in scenarios where a central unit collects information from dispersed IoT devices, such as drone swarms or coordinated vehicle groups, by prioritizing the overall statistical outcome rather than the granular details of each data stream [39].

AirComp capitalizes on the inherent superposition property of wireless channels. This property allows multiple devices to transmit simultaneously over the same time-frequency block, enabling the receiver to directly compute functions (such as the sum or average) of the transmitted signals. By turning interference into an advantage, AirComp enhances spectral efficiency and reduces latency relative to conventional digital processing methods.

Techniques inspired by NOMA further refine this process. In NOMA, overlapping signals from multiple users within the same time-frequency slot are harnessed to improve aggregation efficiency [74]. In addition, recent advancements extend AirComp to relay-based scenarios using a compute-and-forward strategy. In such cases, multiple relay nodes transmit linear combinations of their received messages to a central destination, enabling joint decoding. This relay-based method reinforces AirComp's capability for efficient data aggregation and broadens its applicability in distributed systems.

In recent advancements, the integration of RIS enhances over-the-air computation by dynamically shaping the wireless channel to improve signal alignment at the receiver. By optimizing RIS phase shifts, signal distortion caused by channel fading and interference is mitigated, reducing aggregation errors and improving Mean Squared Error (MSE) performance [75]. This joint design of RIS and AirComp further enhances energy efficiency by minimizing power control complexity at distributed nodes, making it particularly beneficial for large-scale networks with stringent resource constraints. Additionally, RIS can provide an alternative means of compensating for channel mismatches, reducing the need for stringent synchronization requirements that typically challenge AirComp implementations [76].

In distributed systems, where precise individual signal recovery is less critical than extracting meaningful statistical information, in-radio computation offers substantial benefits. It consolidates data from diverse sources at a central receiver, streamlining applications such as Federated Learning, Split Learning, distributed control systems, and tasks like channel and interference estimation. Among these applications, the integration of AirComp within FL has garnered considerable research interest, owing to its capacity to mitigate the substantial communication bottlenecks induced by frequent and large-scale model updates. In this context, several system metrics have been studied to enhance the aggregation and transmission of local models. A significant body of research has focused on minimizing the MSE in the communication channel, particularly by addressing challenges associated with optimal client selection [77], power control [74], and beamforming parameters [78]. Another prominent line of research has concentrated on optimizing the aggregation process, with efforts directed towards the development of channel-aware aggregation strategies, adaptive aggregation algorithms, and methods to reduce the adverse effects of interference and noise on model convergence [79], [80].

While extensive research has optimized the radio aspects of AirComp, reducing the MSE through refined power control, beamforming, and channel alignment techniques, most studies have focused primarily on the propagation side. In contrast, the integration of computational resources, such as task scheduling and the handling of processing delays at edge devices, has not received equivalent attention. This oversight creates a gap in achieving truly end-to-end efficiency, where both communication and computation are dynamically optimized.

To bridge this gap, **6G-LEADER** adopts a ‘compute-when-communicate’ framework that unifies communication and computation through advanced AI/ML methods. Key innovations include:

- **Joint Optimization of Radio and Compute Resources,** where AI/ML-driven channel prediction algorithms dynamically adjust uplink power control and beamforming strategies. Simultaneously, lightweight AI models at the network edge support efficient task offloading and inference on resource-constrained devices.
- **Semantic-Aware Resource Allocation,** where embedding semantic knowledge into xApps and dApps enables the system to prioritize data based on its relevance to the target task, ensuring that critical information is processed with minimal overhead. This semantic layer allows precise resource management, even in densely populated IoT environments.
- **Integration of Advanced Multiple Access Techniques,** where energy- and spectrum-efficient random NOMA and RSMA schemes flexibly allocate resources in line with traffic demands and QoS targets. In addition, reconfigurable PHY-layer technologies, such as

FAs and RISs, are leveraged to shape the wireless environment, reduce interference, and adhere to EMF exposure guidelines.

- **Incorporation of a Robust Conflict Management System,** given that enhanced AirComp techniques can inadvertently induce interference in neighbouring RAN areas. This system, empowered by semantic reasoning, continuously monitors network conditions to detect and resolve conflicts in real time, ensuring that Wireless for AI services achieve the intended accuracy, latency, and energy efficiency.

The proposed framework aims for MSE minimization. Also, by jointly optimizing radio and compute parameters, it achieves notable reductions in overall latency and power consumption. For instance, the system replaces conventional high-complexity convex optimization methods with robust channel forecasting and power optimization strategies that enable rapid adaptation to shifting network conditions.

A concrete PoC will validate these innovations in a real-world scenario. This PoC will demonstrate the practical integration of semantic-aware resource allocation and joint radio-compute optimization. Key performance indicators, including end-to-end latency, energy consumption, MSE reduction, and the effectiveness of conflict resolution, will be rigorously evaluated to benchmark improvements over conventional approaches.

4.2.2 Random and non-orthogonal multiple access schemes

The following presents an overview of random and non-orthogonal multiple access schemes, emphasizing both their capabilities and inherent limitations in supporting massive connectivity and low-latency requirements. The identified research gaps are discussed, followed by the advancements introduced within **6G-LEADER** to enable scalable, efficient, and AI-native multiple access solutions.

4.2.2.1 Scheduled Access

As 6G networks evolve to support vast numbers of connections and extremely low-latency services, traditional orthogonal multiple access schemes are proving insufficient. In conventional approaches, resources such as time, frequency, or code are allocated exclusively to individual users, leading to inefficiencies in highly populated IoT environments and when catering to diverse user needs.

In this context, NOMA emerges as a promising technology for enabling massive connectivity. It operates on the principle of non-orthogonality, allowing multiple users to transmit data simultaneously over the same radio Resource Block (RB) while being distinguished in the power or code domain [81], ensuring efficient data recovery at the receiver. Among the various NOMA techniques, power-domain NOMA is the most widely adopted, as it effectively utilizes power and channel gain differences to multiplex users. Successive Interference Cancellation (SIC) is then applied to the receivers for multi-user detection and decoding. Typically, NOMA considers two user clusters that are sufficiently distinguished in their channel gains (e.g., pairing users with weak channel gains with those with strong channel gains), while research studies user clustering [82], subchannel allocation, and power control [83] to increase resource utilization while balancing system performance and decoding complexity.

More recently, Rate-Splitting Multiple Access (RSMA) has been recognized as a promising multiple access technique in the direction of overcoming the limiting factors of its predecessor power-domain NOMA, which are related to signal decoding complexity and interference management [84]. In the downlink RSMA, the message transmitted to multiple users is split into a common message and a private message. The common message is intended for and decoded by all the involved users in the transmission, whereas the private message is intended for each user separately. As a result, when decoding the private message, the interference stemming from the other users' private messages is treated as noise. In this context, various optimization problems are actively studied, including optimal message splitting into common and private parts, optimal decoding order to ensure effective SIC at the receiver [85], as well as power and rate control strategies. By intelligently and flexibly controlling these parameters, RSMA can strike a good balance between efficient spectrum usage, interference management, user fairness, and signal processing complexity, ameliorating the system's performance.

4.2.2.2 Random Access

Although NOMA enables massive connectivity by accommodating multiple users within a single RB, an additional challenge lies in how each device accesses channel resources. In existing wireless networks, devices request transmission slots through a contention-based random-access process, which introduces significant performance bottlenecks, excessive delays, and signalling overhead. Given the sporadic nature of massive Machine-Type Communication (mMTC) traffic, a gradual shift toward grant-free, contention-based communication is inevitable—allowing devices to transmit data as needed without undergoing the traditional random-access process or by integrating random access with data transmission. Nevertheless, since transmissions occur randomly, there is an increased risk of collisions and interference, potentially degrading performance if not properly managed.

In this context, grant-free contention-based transmission combined with NOMA emerges as a promising solution to enhance efficiency and reduce latency. The primary challenge in power-domain NOMA techniques lies in maintaining an appropriate power difference among users, particularly due to the lack of closed-loop power control. Power-domain NOMA-based uplink (UL) grant-free schemes have been proposed in the literature, such as integrating ALOHA, slotted-ALOHA, or framed slotted ALOHA [86] protocols with power-domain NOMA. In these schemes, the base station dynamically estimates the number of active devices, while novel power control procedures are applied to the transmitters to autonomously select distinct power levels.

Another approach to reducing delay and signalling overhead in sporadic communication is semi-grant-free NOMA [87]. Based on this scheme, grant-free users opportunistically access the spectrum of grant-based users using power-domain NOMA without executing the handshaking process, provided that the latter's Quality of Service (QoS) requirements are met. Effective grant-free user scheduling and optimal decoding order are essential to maximize spectral efficiency, mitigate inter-user interference, and ensure fair resource allocation.

4.2.2.3 Challenges and Research Gaps

While significant strides have been made in NOMA schemes, several challenges persist, particularly as the number of devices continues to grow. Interference management remains a major concern, as accurately decoding overlapping signals becomes increasingly difficult, especially under rapidly changing channel conditions. Moreover, the dynamic allocation of power, rate, and other parameters should not only enhance spectral and energy efficiency but also rely on algorithms that are both robust and computationally efficient to ensure reliable and scalable network performance.

In the realm of mMTC, conventional optimization methods can be overly complex, thereby hindering the system's ability to support a large number of simultaneous connections. This challenge is exacerbated in URLLC, where stringent delay and reliability requirements must be met. Finally, the limited work on random multiple access presents new opportunities for exploration. Metrics such as the outage probability and packet error rate should be jointly considered for evaluating and optimizing the performance of the designed random-access schemes.

4.2.2.4 Building Upon and Advancing Existing Work

6G-LEADER will leverage the advancements in the NOMA, grant-free, and semi-grant-free NOMA schemes to design and propose innovative resource management and optimization solutions aimed at enhancing spectral and energy efficiency. To address the challenges and bridge research gaps, **6G-LEADER** will integrate advanced AI/ML techniques and reconfigurable PHY-layer technologies into the multiple-access design. Specifically, **6G-LEADER** will:

- Utilize AI/ML techniques for channel estimation, active user detection, and traffic prediction in real time, with predictions informing dynamic power control and rate allocation strategies that adapt to fluctuating network conditions, thereby minimizing interference and maximizing both spectral and energy efficiency.
- Utilize AI/ML techniques for user scheduling to shared/non-orthogonally allocated resources, addressing both scheduled and random-access scenarios.

- Utilize AI/ML techniques for solving highly non-convex resource management and optimization problems in general.
- Employ a hybrid approach that combines NOMA with random access, enabling devices to transmit data without prior coordination while still benefiting from interference cancellation methods.
- Leverage innovative reconfigurable PHY-layer technologies, incorporating cutting-edge elements such as FAs and RISs, to actively shape the wireless propagation environment. In this way, adaptive beamforming and effective interference suppression will be achieved, further enhancing the performance of the designed multiple access schemes.

4.3 Goal-Oriented Semantics-Aware Communications

Goal-oriented semantic-aware communication represents a paradigm shift from traditional data transmission, which treats information as raw, content-agnostic payload, or even as sequences of random bits to a model where the *significance, usefulness, and timeliness* of information drive communication decisions. Instead of focusing on delivering all data with high fidelity, semantic-aware approaches prioritize the transmission of task-relevant content that directly contributes to achieving system goals. This reduces unnecessary data exchange, lowers communication, control, and computational overhead, and enhances energy efficiency. These are key requirements for emerging 6G networks and applications such as industrial automation, remote control, and autonomous systems. By integrating AI/ML techniques, semantic-aware communication enables networks to dynamically adapt to context, optimize resources, and support ultra-reliable low-latency performance in increasingly complex and data-intensive environments.

Beyond efficiency, goal-oriented semantic-aware communication introduces new challenges and opportunities for network design. A central question is how to quantify the trade-off between reduced data transmission and the accuracy or reliability of the reconstructed information at the receiver. Metrics such as Age of Information (AoI), Value of Information (VoI), and related indicators have been proposed to capture not just the timeliness but also the utility of information in decision-making contexts. Furthermore, advances in AI/ML enable semantic extraction, reasoning, and representation learning, allowing communication systems to move from raw data exchange toward context-aware decision support. This evolution paves the way for intent-driven networking, where autonomous agents can interpret objectives, dynamically allocate resources, and adapt protocols in real time, ultimately creating communication infrastructures that are more intelligent, resilient, and aligned with application goals.

4.3.1 AI/ML for goal-oriented semantic data networking

With the growing adoption of remote controlled and autonomous systems, such as industrial robots, ensuring seamless operation, real-time monitoring, and responsive feedback mechanisms is becoming more critical than ever. To support the reliable performance of these systems, it is essential to develop realistic testing and optimization scenarios tailored to remote control architectures.

Digital twins are poised to play a central role in industrial automation by reducing operational costs and enhancing productivity. Industrial automation is also emerging as a key vertical in the development of 6G, frequently highlighted in discussions on URLLC and AI-driven edge computing. Achieving effective automation demands a deep understanding of communication networks and their impact on the application context (e.g., robotic behaviour), particularly in scenarios demanding high-precision remote control or when robots rely on edge and cloud computing resources. Optimizing the use of these computational resources is essential for ensuring efficiency and reliability.

Semantic communication emerges as a key enabler in such systems. By leveraging AI/ML techniques, remote operators and autonomous agents (e.g., robots) can exchange high-level, task-relevant information, or even informed decisions rather than raw sensor or video data. This approach significantly reduces data volume while preserving essential context, dramatically lowering bandwidth requirements compared to traditional methods like raw video streaming. A central *challenge* lies in quantifying the trade-off between reduced bandwidth consumption and the system's ability to accurately reconstruct and act upon the transmitted information. In monitoring tasks, a fundamental question arises: how can the project generates a scene description that enables an AI system or an agent on the receiving end to understand and interpret the salient elements of the scene effectively?

Several AI/ML techniques are being considered in goal-oriented semantic data networking, focusing on enhancing effectiveness while minimizing unnecessary information exchange and communication and computational overhead. Recently, this included the integration of Agentic AI, which could introduce a new paradigm of autonomous, goal-driven network management, where AI agents operate proactively rather than reactively, as well as the integration of foundation models (e.g., LLMs) into various layers of the networks and/or in the robotic agents (physical AI). These AI-driven agents do not just optimize the system autonomously based on given constraints, but they have agency, that is they can dynamically reason, plan, decide, and execute actions to optimize data flow, minimize latency, and enhance efficiency based on semantic context and user objectives. With the integration of AI/ML, semantic data networks may dynamically adapt to user goals, data context, application requirements, and real-time conditions, leading to more intelligent and efficient communication networks.

Within the vast ecosystem of AI/ML techniques and architectures, we present below a representative selection of the most widely discussed and promising approaches in the area of semantic data networking:

- *Reinforcement Learning (RL):* RL is a key enabler in goal-oriented semantic data networking, allowing networks to self-optimize, adapt dynamically, and reduce unnecessary overhead. For example, RL-based agents dynamically allocate bandwidth, power resources, and storage based on semantic importance. RL can also detect rare events (e.g., outliers) and mitigate network anomalies in real-time before they cause major disruptions. For that, techniques such as Deep Q-networks (DQN), Proximal Policy Optimization (PPO), actor-critic, as well as Multi-Agent RL (MARL) for distributed network control and optimization, are identified as relevant.

- *Generative AI (GenAI):* Leveraging GenAI and foundation models (LLMs, SLMs), agents can interact with networks using natural language queries, automating configuration changes. That way, human-readable commands can be interpreted and translated into network policies. Going one step further, the communication/networking protocols and the control message exchanges can be treated as a language, and LLM-empowered techniques will enhance efficiency, interpretability, and adaptability. Moreover, semantically irrelevant or useless content could be synthetically generated using GenAI, reducing the need for unnecessary data transmission. Meanwhile, critical information would be processed and transmitted at a higher quality and priority, ensuring efficient resource allocation and optimized network performance.

- *Trustworthy AI*: AI models catering to mission-critical applications or involving algorithmic decision-making that affects users must be reliable, safe, and ethically aligned with human values [88]. They should operate transparently, minimizing risks to individuals and society. Ensuring that AI actions remain transparent, predictable, and controllable as system complexity increases is of cardinal importance. The trustworthiness of the AI/ML techniques developed in the project will be a central focus and a key priority.

- *Representation learning and semantic reasoning:* Key technical challenges in AI/ML for representation learning and semantic reasoning include learning robust and generalizable embeddings from high-dimensional, noisy data; capturing complex semantic relationships beyond surface-level patterns; ensuring interpretability and explainability of learned representations; integrating symbolic reasoning with neural methods; and maintaining consistency and logical coherence in reasoning over structured and unstructured inputs. Specific focus will be on the Platonic Representation Hypothesis [89], which conjectures that the representation spaces of modern neural networks are converging. A highly relevant problem to explore is related to whether the latent universal representation can be learned and harnessed to translate between representation spaces without any encoders or paired data.
- *AI/ML-driven semantic distillation and alignment:* Developing compact AI models that rival or outperform larger ones is crucial. This involves advancing techniques (e.g., RLHF [90], model distillation) that overcome neural scaling laws, enabling AI performance improvements without a proportional rise in cost, power, and energy consumption.

- *Federated and collaborative learning for decentralized network intelligence:* This will enable on-device learning without centralizing sensitive data, reducing network load and privacy concerns. This will also be relevant for multi-agent collaboration, where AI agents deployed across different network nodes can collaborate to optimize network performance in a decentralized manner.

The expected benefits of AI/ML-empowered semantic data networking include:

- *Enhanced effectiveness:* data is contextually prioritized, ensuring only relevant, significant, or valuable information is shared.
- *Reduced overhead:* AI-driven optimization minimizes redundant transmissions and unnecessary computations.
- *Autonomous decision-making and self-supervised learning:* AI agents can continuously refine their strategies without explicit programming and can self-configure and self-optimize network parameters based on real-time conditions.
- *Self-adaptation:* AI/ML allows networks to dynamically adapt to changing or time-evolving patterns (data, traffic, etc.), application requirements, and user needs.
- *Higher efficiency and lower latency:* AI/ML and predictive models enable faster, more reliable data delivery.

Finally, as AI/ML models become more advanced, goal-oriented semantic data networking will continue evolving towards fully autonomous, intent-driven and surprise-inspired networking. This shift will enable intelligent, adaptive, and highly efficient communication systems capable of optimizing network resources on-the-fly while minimizing overhead.

4.3.2 End-to-end information handling schemes

The traditional view of communication systems is that of an opaque, content-agnostic data pipe carrying data, whose value and usefulness for achieving a goal, have been deliberately set aside. This paradigm, although suitable for conventional communication and existing use cases, is inefficient and inadequate to support the data-intensive and timely communication needs of networked intelligent systems [91], [92], [93]. Goal-oriented semantic communication envisions a radically new communication paradigm that accounts for the semantics (importance and effectiveness) of information being generated, processed, and transmitted. A direct gain is an unprecedented reduction in unnecessary data traffic and the associated required communication, processing, and energy resources. Information is useful when it is fresh and timely; this can be captured by the AoI. The concept AoI was introduced recently to quantify the freshness of our knowledge about the status of a remote system [94]. The attention AoI has been receiving is due to two factors. The first is the sheer novelty brought by AoI in characterising the freshness of information versus, for example, that of the metrics of delay or latency [95], [96]. Second, characterising the freshness of such information is paramount in a wide range of information, communication, and control systems. The work in [97] expands the concept of AoI by introducing the Cost of Update Delay (CoUD) metric to characterise the cost of having stale information at the destination. Moreover, the Value of Information of Update (VoIU) metric was introduced to capture the reduction of CoUD upon reception of an update. In pull-based communications, where the endpoint requests and controls the type of the generated information and its arrival time, in [98], the Query AoI was proposed as a relevant freshness metric in such systems. AoI and its recent variants [99], can be seen as simple, proxy metrics of semantics, and have revealed the suboptimality of separate handling of sampling and communication [100]. Information importance can be associated with the "Value of Information" (VoI) in decision/control theory [101]. The works on very-low-latency ultra-reliable wireless communications for industrial control [102] and also in control theory with connection to communications [103], become relevant. However, they do not consider joint handling of information, and the importance of information and its utilisation are ignored.

4.3.2.1 Data fabric for semantic-based data sharing

The latest advances in AI/ML frameworks have brought the focus back on data management. To fully unlock the capabilities of AI, the data used by ML for training and inferencing must be easily findable, understandable, accessible, and reusable, following the FAIR principles [104]. To address these limitations, the data fabric paradigm was conceived, introducing a novel data infrastructure architecture that provides data consumers with a unified access to heterogeneous data silos [105].

Metadata serves as the foundation for the data fabric, guiding the process of raw data ingestion, connecting diverse data silos, and tailoring the exposure of these integrated data to the consumers requirements. This new data management paradigm abstracts data consumers – including operational applications, AI/ML models, or visualization tools – from the underlying complexities and location of the data sources. In this context, metadata functions as a smart integration layer, facilitating seamless data interoperability between various data sources and consumers, as depicted in Figure 4.1. To realize the data fabric, the knowledge graph is envisioned as an enabling technology to build an integration layer based on the semantics of the data and the flexibility of graph structures [106], [107], [108].

Knowledge graphs elevate metadata management and integration, facilitating the seamless integration of disparate data silos. Knowledge graphs unlock the creation of a metadata-driven layer, grounding semantic data models such as ontologies and taxonomies, which capture the concepts and the relations that underpin data silos within the organization[109]. In this respect, standards from the Semantic Web such as Resource Description Framework (RDF), Web Ontology Language (OWL), or (SPARQL Protocol and RDF Query Language) SPARQL [110], are leveraged in combination with existing open-source projects such as Chimera [111].

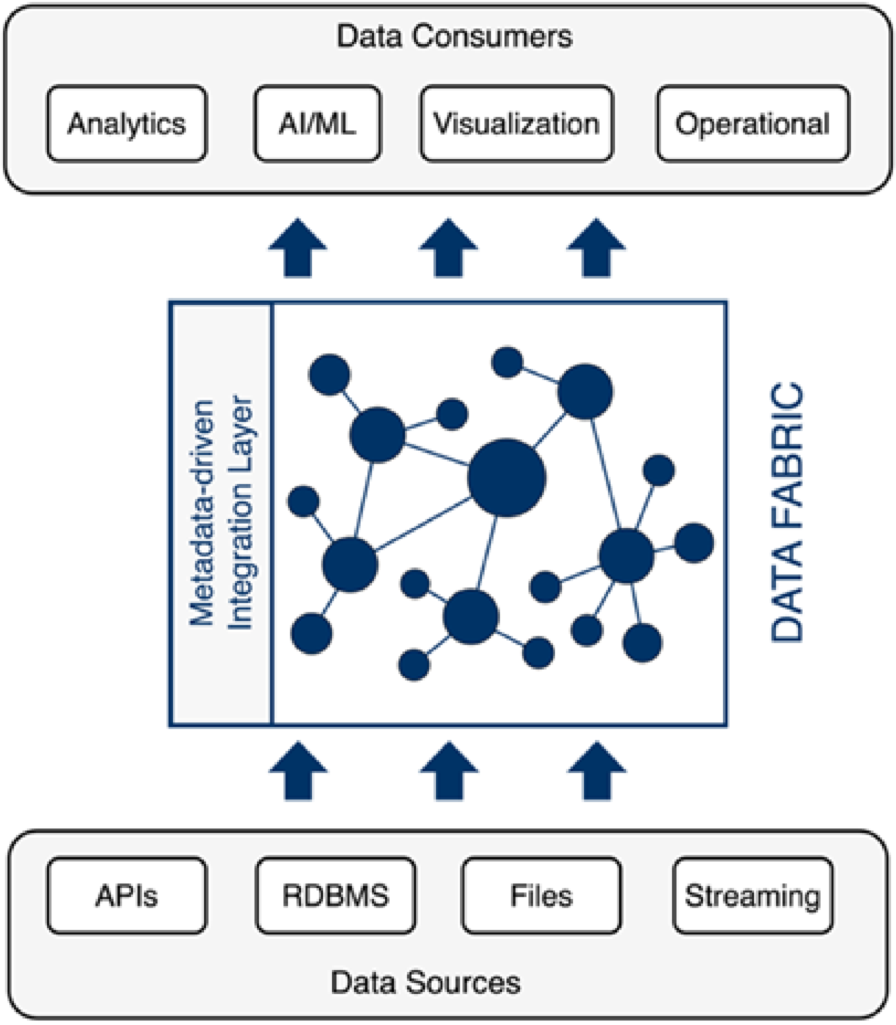

Figure 4.1. Conceptual architecture of the data fabric.

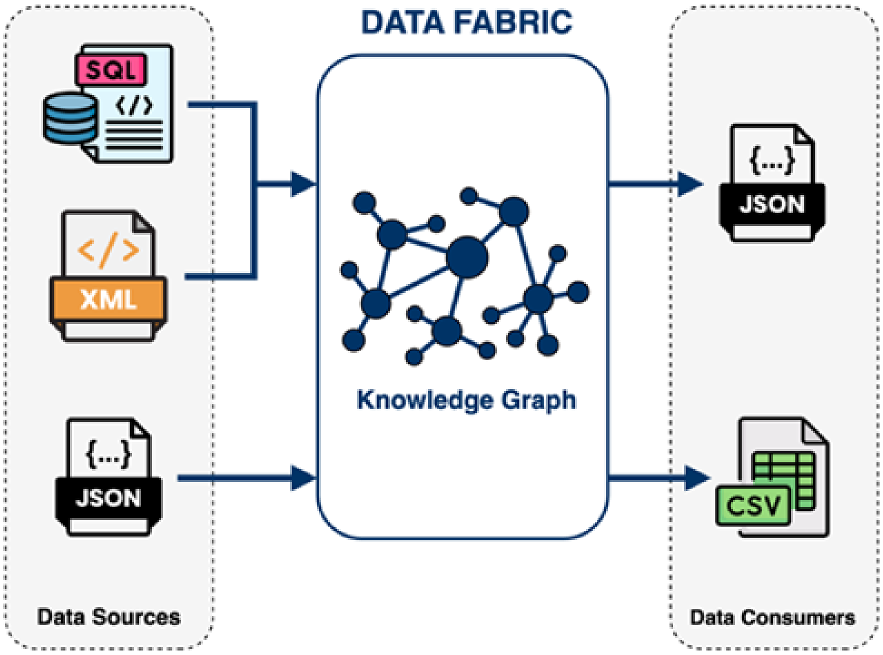

Figure 4.2. Data exchange between heterogenous sources and consumers.

By building upon a knowledge graph, the data fabric enables data sharing by integrating data silos using the proper access protocols and schemas and formats and delivering the integrated to the consumers based on their needs (i.e., destination protocol and data schemas and formats). Figure 4.2 shows a scenario with consumers that requires combined data from a Structured Query Language (SQL) database and an XML file, which in turn, are transformed and delivered as a JSON file.

In this workflow, the data fabric implements specific connectors to extract data from the different data sources. After this step, other components map these data to domain-specific ontologies, transforming and integrating the data into a knowledge graph. Finally, exposure components convert this graph representation into the requested JSON format and schema and deliver it as a file. This same workflow is followed for other use cases, such as transforming JSON data obtained from a message queue like Apache Kafka into a CSV file that follows the structure indicated by the consumer (e.g., ML model).

4.3.3 Energy-efficient semantics-aware schemes

Semantics-aware schemes focus on the meaning and context of the data, rather than just the raw data itself to deliver data communication. Semantics-aware schemes for RAN and O-RAN have been identified as a key technology to overcome the challenges of the future of mobile networks, as move towards 6G [112], [113]. In this context energy efficiency is a key requirement and any proposed solution or technology should be able to address it.

Early research has shown that spectral and energy efficiency can benefit from the integration of semantic communications with O-RAN, and the RAN in general.

The operations of semantic extraction, reconstruction, and classification are important tasks of any semantic communication scheme and can be performed by AI-based autoencoders and classifiers. By emphasizing the transmission of meaningful content over raw data, it is possible to reduce the amount of data that needs to be sent, thereby saving energy and improving efficiency [114]. Techniques such as semantic compression and semantic-aware encoding can be employed to this aim.

Semantic compression schemes in RAN focus on reducing the amount of data transmitted by prioritizing the transmission of meaningful content. This means focusing on the semantic content of the information and transmitting only the essential parts of the data that carry the most meaning. In RAN, these can be used to prioritize critical data, such as control signals and high-priority user data, while compressing less important information.

Semantics-to-Signal Scalable Compression combines semantic and conventional compression techniques. It uses scalable compression to ensure that partial bitstreams are decodable to achieve a certain task (e.g. for machine vision tasks), while the entire bitstream is decodable if a complete reconstruction is needed. This approach allows to minimise the bandwidth usage by transmitting only the necessary semantic information for machine processing, while still ensuring the full data retrieval when needed. Compression parameters can be dynamically adjusted based on real-time network conditions and application requirements, ensuring efficient resource use and energy savings.

Other approaches focus on dynamically adjusting transmission power and data rates based on the semantic relevance of the information [115]. AI and machine learning models can be employed to perform semantic extraction. In this case, the semantic communication is modelled as an optimization problem, aiming to minimize energy consumption while meeting constraints like latency and quality of service.

Hybrid semantic-conventional communication schemes are designed to optimize resource efficiency in sustainable 6G RAN operations by combining the strengths of both semantic and conventional communication methods. This means using semantic communication for applications where context and meaning are more important, and conventional communication for applications requiring high data fidelity and reliability and dynamically switching between the two based on the application's requirements and network conditions. For example, during periods of high traffic, semantic communication can be used to reduce data load, while conventional communication can ensure reliability for critical applications.

Network tomography (NT) is a powerful technique used to infer the internal characteristics of a network by analysing data from its endpoints, such as identifying congested links, detecting faults, and understanding traffic patterns. This detailed understanding of the network's internal state helps in optimizing routing decisions, load balancing, and fault management. This results in network resources used more effectively, potential issues addressed proactively, and configurations adjusted to reduce energy consumption. The insights gained from network tomography can be used to develop semantic models that describe the network's behaviour and performance. These models help in understanding the context and meaning of network data, which is essential for making informed decisions about resource allocation and management.

Network Tomography refers to estimating unobserved network performance metrics from indirect partial measurements obtained from a limited subset of accessible network elements (e.g., nodes or links). It utilizes a subset of monitoring data, corresponding to a partial view of the network state, to perform fine-grained network inference. As such, it is a typical example of an ill-posed inverse problem, where the goal is to determine the underlying factors that produce a set of observations. Such factors may include link-level quality of service parameters (e.g., loss rate, delay, jitter, radio interference), traffic volumes between every pair of nodes in the network (i.e., the Origin–Destination traffic matrix), or the network topology [116]. NT enables efficient network monitoring and presents benefits over traditional monitoring techniques that rely on directly measuring and observing all elements of interest. Specifically, it reduces computational and traffic overhead compared to other packet-level, flow-level, and signal sensing monitoring methods while alleviating the need for explicit cooperation and participation of all network elements, which improves scalability [117]. However, a trade-off between overhead reduction and estimation accuracy must be carefully considered depending on application requirements.

In mobile networks, spectral efficiency has increased through advanced multiplexing strategies that are coordinated by base stations (BS) in licensed spectrum. However, external interference on clients leads to significant performance degradation during dynamic (unlicensed) spectrum access (DSA) [118]. As spectrum sharing moves towards lightly licensed and unlicensed models, DSA continues to be an important issue for better use of our critical spectral resources. Issues such as the hidden terminal problem can be decisive for system performance and the successful development of Access-Aware (AA) schemes, which incorporate some knowledge of the interference, e.g., namely the probability that individual clients can access the channel, in the scheduling and resource allocation decisions. However, estimating interference is, in fact, a (receiver) location-dependent one. Thus, even a sophisticated spectrum scanning solution located at the BS cannot obtain a comprehensive view of the interference environment.

At the same time, the traditional wireless connectivity paradigm of neglecting the context-dependent meaning of transferred data is shifting towards approaches that make the semantics of information [91], i.e., the significance and usefulness of messages, the foundation of the communication process. This is unavoidable, as cyber-physical and autonomous networked systems handle large sums of distributed real-time data that end up being useless to the end user and causing communication bottlenecks, increased latency, and safety issues. This entails a goal-oriented unification of information generation, transmission, and reconstruction, by considering process dynamics, signal sparsity, data correlation, and semantic information attributes.

Network tomography can be used to address such challenges in DSA systems. More specifically, **6G-LEADER** will explore the use of network tomography techniques for DSA in two main directions: a) towards accurate interference and channel state information feedback, which will be exploited for better resource allocation, and b) towards semantics and context awareness that will be exploited in scheduling decisions. More specifically, in the first direction, the project will adapt AI/ML-based tomographic approaches for channel estimation and interference inference in a local radio environment with minimum user-provided signal sensing information [119]. Utilizing this information can lead to a more effective orchestration layer, optimizing network and computing resource usage across complex and heterogeneous transmission conditions. Along the second direction, the project will extend and further develop AI/ML-based NT frameworks [120], [121] to account for traffic context and network load when providing orchestration information, facilitating proper adaptation to the semantics of transferred data and efficient optimization of the involved network parameters.

Overall, NT enhances network observability without increasing the volume of measured data, thereby reducing the required monitoring demands. Leveraging existing information to generate insightful estimates improves efficiency, lowers equipment and operational costs, and facilitates verification of service-level agreements.

4.3.3.1 Challenges and Research Gaps

Energy-efficient semantics-aware schemes for RAN and O-RAN are crucial for the future of mobile networks, especially as we move towards 6G, but they come with some challenges:

Energy consumption in mobile networks is a significant concern, particularly regarding RAN. The increase of energy usage not only leads to higher operational costs for the Mobile Network Operators (MNOs) but also has a clear environmental impact. Leveraging on AI/ML to optimize the power usage can be an option only if the energy-efficiency of the algorithms themselves are considered. How to optimize for semantic-aware schemes for energy efficiency is still an open question.

Early research has shown that spectral and energy efficiency can benefit from the integration of semantic communications with O-RAN, and the RAN in general. However, this integration can be complex and may require sophisticated tools for data processing and knowledge extraction. Therefore, there are still many aspects to be investigated to make this integration effective.

How to evaluate the sustainability of semantic-aware schemes and which metrics to use is still an open question for which there is still room for discussion. Since the field is still young, there do not exist consolidated best practices on who to integrate semantic communications into existing RAN architecture.

4.3.3.2 Solutions proposed in 6G-LEADER:

6G-LEADER will leverage Network Tomography (NT) to reduce monitoring overhead and provide accurate input for network management tasks such as resource allocation and orchestration. Specifically, since the linear measurement model in NT is directly analogous to the linear model used in channel estimation, advanced ML-based methods originally developed for NT (e.g., deep generative models such as Variational Autoencoders (VAEs), Generative Adversarial Networks (GANs), and Invertible Neural Networks (INNs)) will be adapted to reconstruct channel state information (CSI) and potentially other hidden network parameters from limited measurements. By enabling accurate and efficient CSI with minimal sensing, NT will support and facilitate semantic communications, where reliable CSI is crucial for carrier selection, allocation between semantic and conventional streams, and balancing spectral efficiency, energy consumption, and semantic reliability. In addition, NT-derived insights (e.g., estimated traffic load) will be integrated into the developed intent lifecycle management framework described in the next section to enhance validation, assurance, and adaptive decision-making regarding orchestration actions, thereby contributing to the efficient and sustainable operation of 6G systems.

4.3.4 AI-driven Intent Lifecycle Management

Intent, as outlined in IETF RFC 9315, refers to a declarative specification of desired operational goals and outcomes, without prescribing the methods for achieving them. In essence, it represents a high-level expression of constraints and optimization objectives that need to be met during the deployment and operation of networked services or applications. Depending on the service delivery model and the roles of various stakeholders, e.g., vertical application providers, infrastructure operators, or communication service providers, the responsibility for defining intents may vary. Regardless of who defines them, enabling effective and intelligent intent lifecycle management introduces several key challenges that must be addressed.

Intent-Based Networking (IBN) is increasingly recognized as a fundamental enabler for autonomous service and network orchestration in 6G environments. Central to IBN is the definition and implementation of an intent lifecycle management module that ensures the continuous satisfaction of the intents deployed on the system. Apart from Intent Representation, Intent Translation, Policy Mapping and Intent Verification stages, an IBN should also implement the Intent Assurance stage which leverages monitoring data from the deployed intents and the infrastructure layer in order to report that status of the intent back to the user, provide performance assessment, and take corrective action towards intent satisfaction [122]. This assurance module should be implemented across all layers of the 6G platform, encompassing the Radio Access Network, Transport Network, and Core Network.

Building on this foundation, a principal line of research has focused on developing end-to-end platforms that translate natural language intents into actionable network policies using deep reinforcement learning, while continuously adapting configurations across multi-domain networks through real-time monitoring data [123]. These platforms incorporate modules for natural language processing, log analysis from sources such as Prometheus and Elasticsearch, and policy generation, which is executed via orchestration control layers. Another key research direction addresses the challenge of intent conflict—where the satisfaction of one intent may impede the satisfaction of another—through closed-loop optimization frameworks. These approaches are based on game-theoretic models, such as the Weighted Nash Bargaining Solution (WNBS), the Kalai-Smorodinsky Bargaining Solution (KSBS), and the Shannon Entropy Bargaining Solution (SEBS), to effectively detect and resolve conflicts [124].

Complementary efforts investigate the use of Large Language Models (LLMs) to express, refine, and validate intents in natural language. Quantized low-rank adapters are used for fine-tuning LLMs to enhance resource efficiency. Furthermore, transformer-based forecasting mechanisms, such as Retrieval Augmented Generation (RAG) and the Informer model, are utilized to predict network conditions, including traffic load and power consumption. Finally, a Hierarchical Decision Transformer with Goal Awareness (HDTGA) has been proposed to guide orchestration decisions and optimize overall network performance [125]. The integration of LLM for processing multimodal intents (i.e., intents expressed in natural language accompanied by deployment descriptors) has also been investigated. The LLM provides a template describing the optimal deployment policy, which is then converted into a deployment-ready service order in a standardized format [126].

Regarding the O-RAN platform, one approach is to design an IBN system with a closed-loop architecture, where an Event Calculus logic model is employed for intent goal modelling and further goal decomposition and reasoning. Based on a continuously updated Knowledge Base, the resource allocation problem is formulated as a Markov Decision Process and addressed through a deep Q-networks algorithm, which produces new rules/policies for initial deployments and corrective actions, while updating the Knowledge Base [127]. Finally, there has been a remarkable effort to standardize the intent lifecycle architecture for multi-tenant 6G scenarios, defining the architecture layers, interactions, and responsibilities of each stakeholder [128].

Within the context of **6G-LEADER**, it will adapt and extend an Intent-Lifecycle-Management (ILM) framework for managing distributed services represented as annotated graphs of components and links [129]. The proposed architecture operates through three nested control loops spanning the user, processing, and implementation spaces. In the first control loop, high-level intents comprising objectives and constraints (expressed in the User space) are semantically validated to detect conflicts or infeasible goals. These validated intents are translated into machine-readable deployment plans using natural language processing and a TOSCA-based descriptor [130]. An optimization solver generates proposed plans, which can be stored in a shared knowledge base for further analysis and review. After the successful deployment of the application on the infrastructure nodes, runtime metrics are continuously monitored and leveraged by the second control loop to detect or predict intent violations and trigger quick, short-term adaptive actions. Finally, a long-term intent fulfilment report is generated by the third control loop, which informs the user about the execution of their intent and suggests refinements to the intent parameters.

To advance beyond the current state of the art and enable more context-aware and efficient interpretation while minimizing resource consumption, the project will enhance the integration of LLMs with structured knowledge bases, such as knowledge graphs, for both intent expression and processing. Reinforcement learning (particularly hierarchical and agent-based AI techniques) will be explored to support adaptive and long-term policy optimization across diverse network environments. Service profiling, combining historical data with real-time measurements, will be used to strengthen intent validation by aligning it with empirically observed behavioural patterns. Forecasting techniques will play a greater role in estimating the likely behaviour of intents within the system, supporting more informed decision-making. The consortium will also expand the range of orchestration actions available to the management system. Finally, the project will leverage LLMs not only for lifecycle management but also for natural language error reporting and, potentially, for intent retrieval directly from the system state.

5 Reconfigurable Components and O-RAN Functionalities

As next-generation communication networks evolve to meet the demands of new applications and services, the integration of reconfigurable RF components has emerged as a potential direction towards more flexibility, higher efficiency as well as better performance. Two promising technologies in this area are FAs and RIS. FAs offer dynamic adaptability in terms of their shape and position, enabling real-time configuration and optimization of the transmitted/received signal, while RIS can control the propagation environment to enhance the coverage, the energy efficiency and the signal strength through software-controlled surfaces. Apart from these hardware-based technologies, the Open Radio Access Network (O-RAN) architecture introduces an open, interoperable and programmable interface that can support the integration and coordination of various network components, including FAs and RIS. All these technologies represent a shift towards more adaptive, reconfigurable and software-based wireless networks. A state-of-the-art analysis as well as a discussion on the innovations of **6G-LEADER** are presented in the following sub-sections.

5.1 Reconfigurable RF Components

Reconfigurable RF components can reversibly change and adapt either their physical shape (in the case of FAs) or the signal's electromagnetic properties (in the case of RIS) to match different specifications or requirements. Several advantages naturally arise from this feature, as they improve communication capabilities when compared to traditional, static systems by supporting greater flexibility and responsiveness in wireless environments. In general, they unlock additional degrees of freedom in hardware and signal processing design. Traditionally, the application of reconfiguration mechanisms has been focused on radiation pattern and frequency modifications, as they are closely related to the physical dimensions and structure of the antenna. Consequently, they are particularly interesting to explore novel frequency ranges such as FR1/FR3, of relevance for this network.

5.1.1 Fluid antennas

Future wireless networks will embrace numerous technologies and devices within a reliable and systematic architecture of applications and services. To be able to adapt to the continuous variations in demands but also in the physical environment, it is essential for these networks to be reconfigurable and intelligent. From a transceiver's point-of-view, this adaptability can be accomplished through the employment of reconfigurable FAs. These refer to antennas that are flexible in the sense that they can alter their physical structures and/or adjust their electrical characteristics to support different configurations, for example, with respect to the operating frequency, radiation pattern, and polarization. This reconfiguration can be achieved through a programmable and controllable manner.

Multiple-input multiple-output (MIMO) systems have been an integral component of wireless communication systems since the introduction of the 3G technology. Theoretically, MIMO provides a throughput increase proportional to the minimum number of transmit and receive antennas. Nevertheless, many antennas corresponds to complex RF signal processing but also requires sufficient spacing between the antenna elements to mitigate mutual coupling. Both limitations increase the size of the antenna array, their implementation cost, as well as their power consumption. In view of this, the FA technology provides new degrees of freedom in the design of wireless communication systems and has the potential to address fundamental design restrictions and push further the performance limits of wireless networks. Indeed, FAs can assist with various network demands for higher data rates, interference management, higher reliability, and energy efficiency. Due to the flexibility and performance gains achieved by FAs, compared to conventional antennas, there have been significant research efforts recently towards their exploitation and further utilization in wireless networks.

Most of the literature on FAs consider a tube-like linear architecture within which the liquid is moved. Specifically, a microfluidic system can alter the location of the liquid to one of the preset locations, also known as “ports”, that are evenly distributed along a linear dimension. Therefore, the shape of the FA cannot be changed but its position can be adjusted to extract diversity and multiplexing gains. Recent works investigate the concept of FAs in the context of point-to-point communication systems, where a mechanically flexible single-element antenna over a small linear space is employed, and the achieved performance in terms of outage and ergodic capacity is evaluated as described in Figure 5.1 [131], [132]. A key finding is that though space matters, a single-element FA with a tiny separation of half-wavelength or less between the ports can deliver capacity and outage probability that is achieved by a multi-antenna Maximum Ratio Combining (MRC) system, if the number of ports is large enough. Within the framework of multi-user communications, a mathematical framework has been developed that takes into account the existence of multiple pairs of transmitters and receivers, whereas a selection combining technique has been adopted at the receivers to switch their single-element linear FA to the position with the strongest SIR [132]. Exact and approximated expressions for the outage probability, capacity, as well as multiplexing gains have been obtained, illustrating that the network multiplexing gain grows linearly with the number of ports at each receiver while it is ultimately limited by the number of receivers. Furthermore, the capability of such communication systems to support hundreds of users by using only one FA at each user is illustrated, giving rise to significant enhancement in the network outage capacity.

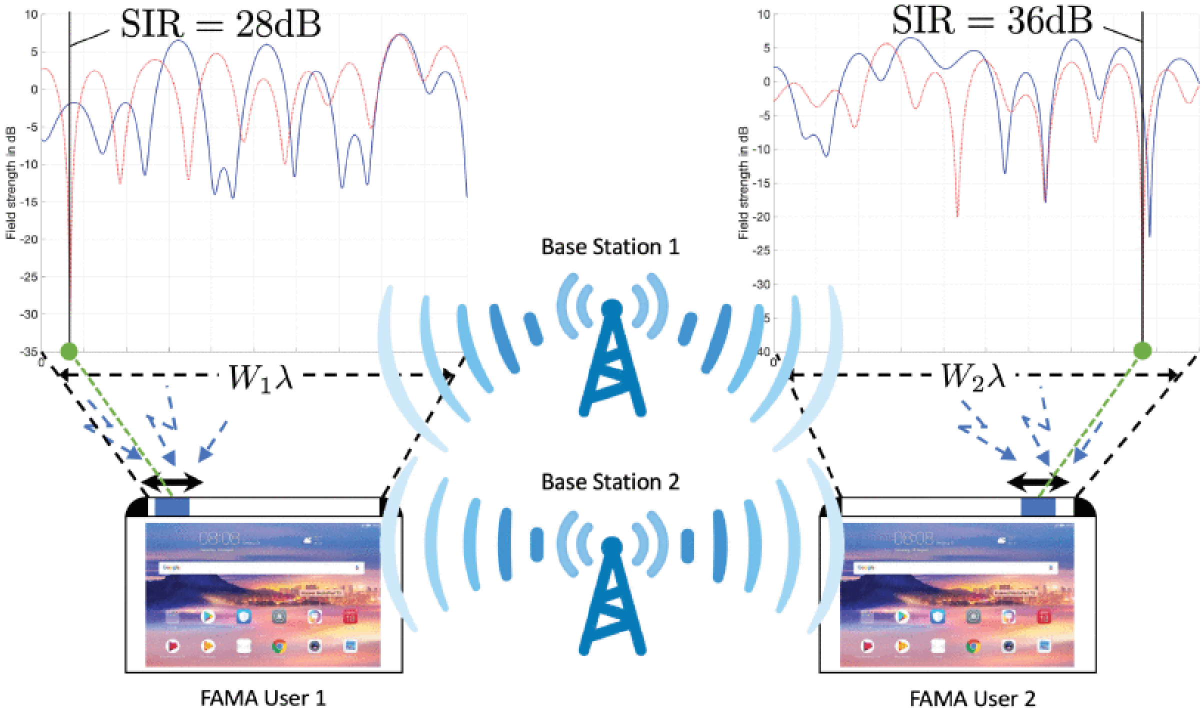

Figure 5.1. Fluid antenna multiple access (FAMA) scheme [132].

Well understood state of the art communications techniques such as MIMO, beamforming or multiple access have been complemented using FA. But also, other emerging technologies such as NOMA or Integrated Sensing and Communication (ISAC) are being explored for FA. Besides theoretical analysis of FA, some prototypes can be found in the literature as well. Of particular interest for this project are the ones based on liquid metals, which usually employ eGaIn (eutectic alloy of Gallium and Indium) or Galinstan (adding Tin). These Gallium-based alloys are liquid at room temperature and biocompatible, that is, they are not toxic, radioactive nor flammable unlike their traditional competitors. Additionally, they present good electric properties that make them suitable for RF applications.

Fundamentals of FA implementation can be seen in [131]. There, a Yagi-Uda antenna is designed using liquid metal to vary the height of a column of metal in different tubular deposits. In consequence, antenna dimensions are variable with time, achieving reconfigurability. Undoubtedly, having several syringes is impractical to control the flow of liquid metal in commercial applications, but it represents a proof of concept for implementations of liquid antennas.

Indeed, mechanical means like syringes are likely to degrade with use due to their movable parts. Hence, it would be desirable to displace the metal by applying electrical impulses only. This is possible thanks to a technique called electrowetting. It allows varying the rheological properties of the material in a controlled manner[48], which can be employed to control movement. Therefore, analogous to transistor biasing processes, DC signals can be utilized to control the motion of the drop while RF signals are employed concurrently to carry out data transfer. Note that understanding port distribution is crucial to FA designs. Generally, ports are continuous throughout the device, which means that there is not "empty" space between them.

One parameter of key importance for the performance of liquid FAs that is directly addressed in this project is reconfiguration speed. While new generation technologies tend to increase the carrier frequency, reconfiguration speed becomes more challenging during the implementation stage. Alternative designs aim specifically to address this problem, such as reconfigurable pixel antennas or mechanical movable antennas. However, these options do not fully exploit fluidic properties of liquid metals. Although reconfiguration is not possible at symbol rate, liquid-based FAs still have suitable use cases, as presented above.

Other solutions employ liquid metal as reflectors or directors [133] that dictate the radiation pattern of the antenna. This solution offers simplicity in the design while working on the FR3 band. Hybrid solutions combine in more or less degree the solutions at the cost of potential increases in design complexity.

Despite the growing interest and the development of several experimental demonstrations, a deep understanding of FA systems is still missing. Indeed, the theoretical and practical limits of their use in real-world wireless communication systems have not been fully established. As such, advanced signal processing methods are needed that can effectively exploit the reconfigurable liquid nature of these antennas to unlock the potential gains in diversity and multiplexing gain. The integration of FAs within modern communication systems presents several challenges. Specifically, FAs introduce non-conventional spatial characteristics that do not follow traditional channel models, thus requiring new mathematical frameworks. Moreover, realizing instant and energy-efficient displacement of the liquid remains an important challenge, especially at higher frequencies relevant to 6G, such as the FR1 and FR3 bands. These bands impose stringent requirements on the positioning precision and reconfiguration latency both of which must be addressed to ensure robust performance. The **6G-LEADER** project aims to address these challenges by developing next-generation physical-layer solutions that fully take advantage of the FA technology. This includes novel beamforming techniques, coding and modulation schemes, and new multiple access methods. All of these will be designed in such a way to exploit the reconfigurable features of an FA. Additionally, the project will demonstrate the viability of FA-enabled systems in the FR1/FR3 bands, where traditional fixed position antennas have significant limitations. By exploiting the additional degrees of freedom offered by FA architectures, **6G-LEADER** aims to establish FAs as a foundational technology for reconfigurable, low-latency, and energy-efficient 6G communications.

5.1.2 Reconfigurable intelligent surfaces

RIS represents a transformative approach in the design of next-generation wireless communication systems. These surfaces, composed of a large array of passive or semi-passive elements, can dynamically manipulate electromagnetic waves to enhance signal propagation, coverage, and energy efficiency. In this subsection, we explore the fundamental structure of RIS, as well as a range of use cases that highlight the versatility of RIS in some specific scenarios. These capabilities position RIS as a key enabler in the evolution of reconfigurable RF components. However, the introduction of RIS as part of the 6G standardisation process is not guaranteed because the deployment of these components in a massive mobile network deployment is not straightforward and probably it is not scalable.

Additionally, the most common assumption is to consider RIS as a new network element between the base station and the user equipment. This scenario contains multiple challenges for MNOs such as optimal RIS placement, regulations, interference management, among others.

On the other side, it is envisioned that 6G will implement extra-large massive MIMO scenarios. It is going to guarantee, among others, that the same 5G grid of cells could be reused for 6G. Increasing the number of antenna elements and MIMO streams results in an unscalable increase in the hardware complexity. One of the solutions to this complexity in 5G was to include not only digital beamforming but also hybrid beamforming to reduce the number of RF chains. However, hybrid beamforming is probably not enough to face the expected increase in the number of antennas and MIMO streams. For this reason, **6G-LEADER** is also considering the integration of the RIS into the radio unit substituting the analog beamforming to drastically reduce complexity and cost.

A basic RIS system is composed of an array of tuneable elements that can be controlled by a RIS controller. The elements are dynamically adjusted to control the reflection coefficient of the surface, steering the desired signal to a specific direction. Figure 5.2 shows a basic scheme of the RIS principle, in which the reflected wave direction depends on the amplitude and phase profile of the RIS elements (also known as RIS configuration).

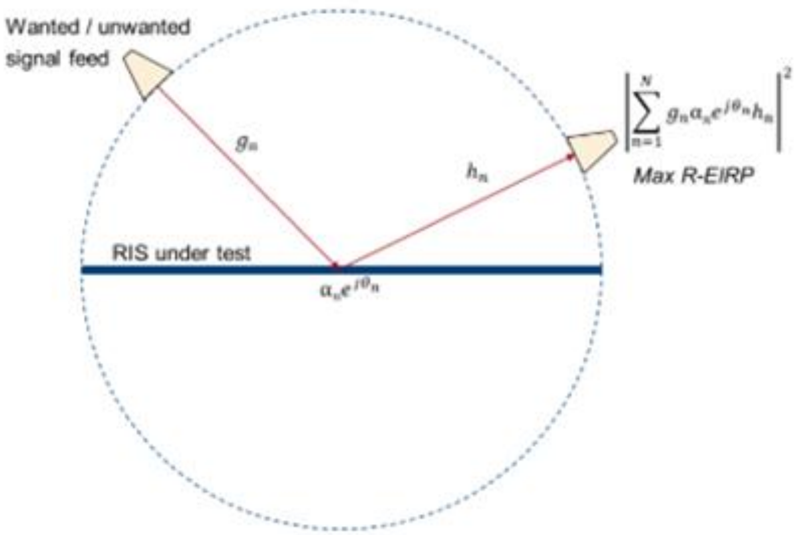

Figure 5.2. Basic RIS principle [134].

The RIS integration in mobile networks, including the O-RAN architecture, is complex and demands consideration of multiple aspects. Especially, the integration depends on the use case and the type of RIS controlling considered. According to [134], the RISs have been classified into different categories considering the type of control.

An example of this category is the network-controlled RIS, in which the network determines the control information that is used by the RIS for control and configuration. The network determines the information based on the collected data from the RIS and/or the UE. Additionally, the RIS can also provide the network information collected from the UE. Consequently, the network should be able to process the information to decide the RIS configuration and communicate it to the RIS controller. A possible solution to support this management and/or control of the RIS should be the definition of new interfaces between the O-RAN entities and the RIS controller.

On the other hand, the integration of the RIS can be conditioned by the considered use case or the RIS topology. ETSI has also exposed different topologies in the context of some RIS use cases [1]. For instance, different RIS topologies for communications, for localization, and for improving ISAC systems with passive or active sensing.

Focusing on the specific use case of communications, RIS can be used in three different topologies: Case A, B, and C.

- Case A: The RIS is co-located with or integrated as part of the transmitter. For instance, the RIS can be used to replace the conventional phase shifter and power amplifier in a Massive MIMO transmitter. This topology consists of only two elements: the RIS-based transmitter and the Receiver.
- Case B: In this topology, the RIS is an intermediate entity between the transmitter and the receiver, which is allocated in distributed locations. In this case, two scenarios are considered: extended coverage in holes of an outdoor scenario and/or weak coverage that can happen in both indoor and outdoor due to blockage. The analysis of this project is mainly dedicated to this case, which is shown in Figure 5.3.
- Case C: However, ETSI also considers Case C, in which the RIS is also allocated in a distributed manner, but it is not limited to 3GPP networks. It can be also used to interconnect customer premises networks, personal internet of things networks, device-to-device communications or Wi-Fi.

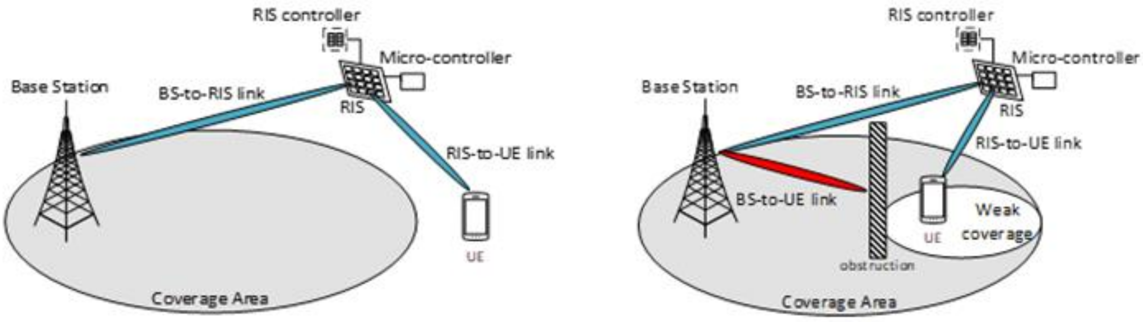

Figure 5.3. Improved coverage a) Outdoor. b) Behind obstacles.

The type of control in the RIS and the different topologies have been presented as fundamental points to consider for network integration. However, they are not the only aspects to consider. Additionally, the specific use case in which the RIS is employed can require particular attention in the standardization process. In this regard, RIS can be seen as an enabler of multiple use cases, in which the enhancement of ISAC systems and localization systems are outstanding use cases, the integration of the RIS in such scenarios is commonly associated with additional challenges such as the proper RIS selection.

The RIS selection process should consider multiple aspects. For example, the number of RIS that are considered, the simplest case is when only one RIS is considered in the path. However, a multi-RIS scenario can be beneficial or required to reach the target area for sensing or localization. In this specific case, the RIS selection procedure will have additional complexity and will demand multiple consideration. Additionally, the frequency band in which the RIS is implemented can have a significant impact on the performance of the whole system. For this reason, it is planned to analyse the performance of the RIS comparing different frequency bands, specifically FR1 and FR3.

Moreover, user scheduling in 6G RAN architecture design is crucial for managing limited radio resources and ensuring fair and optimal performance for all users. This need becomes even more pronounced with the adoption of advanced antenna technologies such as new hybrid beamforming schemes in which the analog beamforming part is replaced by a RIS. While much of the existing literature assumes that all UEs are served in every time slot, in practice, this is infeasible for large-scale systems, therefore, we require algorithms that schedule only a subset of users in each time slot.

In general, scheduling objectives are designed to optimize long-term fairness criteria, such as Proportional Fairness. This involves assigning a scheduling weight to each UE based on its historical service, with the objective of maximizing the aggregate weighted sum rate within each scheduling interval. A key challenge introduced by modern antenna systems and RF components is the complexity of the instantaneous rate vector space across UEs, which complicates the scheduling process.

In earlier work from Nokia on Hybrid Beamforming [135], [136], [137] this problem is addressed using a sequential approach, as illustrated in Figure 5.4. First, an optimal analog beam is selected for each UE, allowing the estimation of the maximum achievable rate for that UE in isolation. Next, users are selected to maximize the weighted sum rate, considering both individual rate estimates and the inter-user interference. For instance, users that are far apart in beam space are more likely to be chosen. Finally, a digital beamforming algorithm is used to compute an optimal precoder for the selected users.

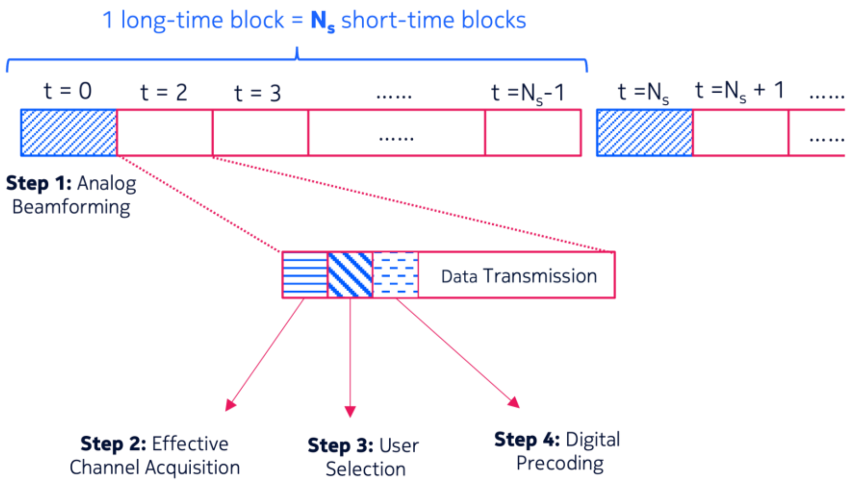

Figure 5.4. Baseline algorithm for user scheduling in hybrid beamforming [135], [136], [137].

This framework should be extended to scenarios involving near-field RISs, FAs and non-orthogonal transmission to address effective evolution of reconfigurable RF components in FR1 and FR3 bands. These extensions are critical for supporting the continued evolution of reconfigurable RF components in FR1 and FR3 frequency bands. In each case the main problem is to estimate the achievable rate vector for a given subset of users based on their spatial geometries and propagation characteristics.

5.2 O-RAN-based Cellular Architecture

This section explores key enhancements to the O-RAN architecture for future 6G networks. Subsection 5.2.1 explores how AI/ML and semantic communication are driving the evolution of intelligent, autonomous O-RAN systems for 6G networks. Then subsection 5.2.2 discusses critical extensions, to both O-RAN components and interfaces, needed to overcome current technical limitations and enable advanced capabilities, such as E2 interface enhancements to support for semantic-empowered xApps and real-time closed-loops custom logics. Finally, the importance of coordination mechanisms to manage conflicts between RAN applications is addressed in subsection 5.2.3.

5.2.1 AI/ML and semantics for O-RAN

The O-RAN is evolving rapidly and becoming more intelligent with the integration of AI/ML and semantic communication technologies. These technologies are helping O-RAN create smarter, more autonomous and more efficient networks. The O-RAN Alliance's Next-Generation Research Group has laid the foundation for what it calls AI-native networks. These networks are designed to use distributed intelligence, digital twins and semantic communication methods to manage resources more effectively and respond to network changes in real time [138], [139].

A recent example of this is O-RANSight-2.0, a domain-specific LLM for O-RAN that uses retrieval-augmented generation (RAG)-based instruction tuning framework with two LLM agents [140]. The proposed framework outperforms general-purpose LLMs like GPT-4o for RAN-related tasks. Moreover, a different approach is a neurosymbolic-based Federated Machine Reasoning (FLMR) method, which is a transparent and effective AI/ML decision making option for dynamic O-RAN systems [141]. It optimizes the CPU demand in virtual base stations and achieves an effective balance between resource overprovisioning and under provisioning. Another area of interest is explainable AI (XAI), which assists towards understanding how AI systems make decisions. This is important for building trust and allowing for better control over automated network functions [142]. For example, the EXPLORA system provides detailed explanations of deep reinforcement learning decisions used in resource management [143]. Also, a new SMO framework was designed to support a centralized ML architecture for training and policy control, to address the demands of managing the complex O-RAN interfaces and components[144], [145]. Other research directions, such as lightweight ML-based xApps for real-time resource control, have also shown good performance in meeting quality-of-service targets in near-real-time RIC environments [146].

Semantic communication is also becoming an integral part of future O-RAN and 6G systems. Instead of just sending raw data, semantic communication focuses on sending the actual meaning or intent behind the data. Recently, researchers have proposed new architectures that include components like a semantic RIC and a semantic plane, which support intelligent decision-making based on context [113]. For example, the SEM-O-RAN uses semantic-aware slicing to improve the offloading of computer vision tasks to the edge. By applying class-based image compression and flexible slicing, SEM-O-RAN can handle up to 169% more tasks without reducing the quality or speed [147]. Moreover, a Semantic-Aware RAN (S-RAN) system offers a holistic solution for semantic communication beyond single transmission pair [148]. Finally, a digital twin-enabled O-RAN architecture with semantic communication has been proposed to support ultra-reliable low-latency communication. This system uses real-time representations of the network to make fast and reliable decisions, which is especially useful in demanding applications like smart manufacturing [149].

Overall, these efforts are helping O-RAN become a smarter and more self-managing network system. With the use of AI/ML and semantic communication, future networks will not only be faster and more efficient but also more flexible, transparent and ready to meet the demands of 6G. Towards achieving this vision, **6G-LEADER** is extending AI/ML capabilities into the O-RAN to support real-time applications with response times under 10 milliseconds. It introduces distributed applications (dApps) that collect real-time data and performance metrics from O-RUs, O-DUs, and O-CUs, while also using additional context from near-real-time RICs to control lower-layer radio functions. The main innovation of **6G-LEADER** is the semantic alignment between these components, which allows smarter decisions to be made closer to the radio layer.

5.2.2 O-RAN extensions

The O-RAN paradigm has provided 5G systems with a significant advance from the traditional RAN approach, by promoting an open, disaggregated, and intelligent architecture. Moving away from a monolithic implementation, O-RAN facilitates the distribution of the RAN functions by dividing the gNB into three main components, namely the Central Unit (CU), Distributed Unit (DU), and Radio Unit (RU) [150], and by defining standard (open) interfaces and different functional splits, such as the 7.2 which relies on the Open Front Haul (OFH) specifications [151]. Disaggregation greatly enhances scalability, flexibility, and vendor interoperability, while facilitating RAN function virtualization (vRAN). Moreover, RAN softwarization and intelligence are at the heart of O-RAN, which has also introduced the RAN Intelligent Controller (RIC) to support the integration of third-party AI-based applications for dynamic network control and optimization. The E2 interface connects the CU and DU to the near-real-time (RT) RIC to deploy xApps and implement control loops, with latencies between 10 milliseconds and 1 second [152]. Similarly, the O1 interface connects the gNB to the non-RT RIC to deploy rApps and implement non-time-sensitive tasks aimed at Service and Management Orchestration (SMO) automation [153]. Management of the RU is possible using the M-plane over the OFH. The O-RAN management capabilities are further expanded by the O2 interface, which connects the O-Cloud with the SMO framework, and the A1 interface which interconnects the non-RT and near-RT RICs. Figure 5.5 shows the different interfaces leveraged for SMO in the O-RAN architecture.

Whilst the paradigm shift promoted by O-RAN serves as a solid base upon which design the future 6G systems, several relevant technical challenges stem from its current architecture. Representatively, semantics-empowered communications, a key element in the design of the 6G RAN (e.g., intelligent resource management), are not considered in the current O-RAN architecture. Although the topic has attracted the attention of the academia in the last years [112], [147], the current O-RAN specifications lack the means to exploit semantic-aware AI-based

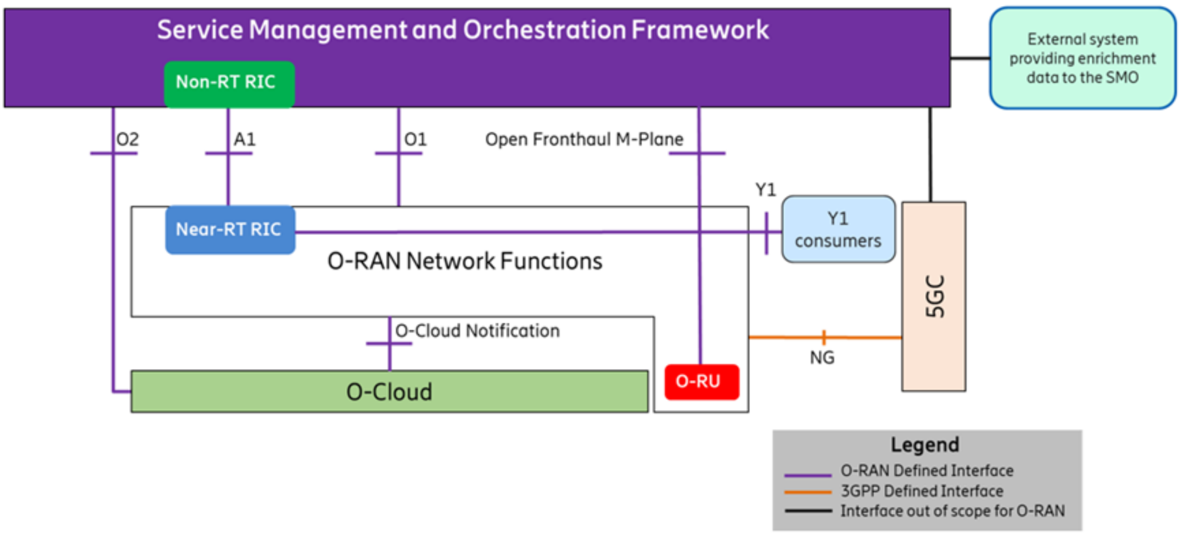

Figure 5.5. High-level O-RAN interface overview [154].

applications, efficiently and in a standardised manner. In this regard, the definition of the E2 interface is not currently considered support for semantic-empowered xApps. Moreover, it offers limited access to the Physical (PHY) and Medium Access Control (MAC) layers of the DU, which in turn constrain the AI-enabled optimization of various RAN aspects that are directly related to key 6G KPIs. For instance, energy efficiency and EMF exposure reduction can be improved through fine-grained control of RIS-based hybrid beamforming and optimum spectrum usage of coexisting FR1 and FR3 bands [155], but this requires access to the channel state information, precoder, and scheduler of the DU [156], which is not currently contemplated. Additionally, the current E2/xApp framework only considers closed loops with latencies of 10 milliseconds or more. Advancing to true real-time closed-loops is currently under analysis by the O-RAN Alliance [157] and the focus of several interesting works in the academia that focus on the definition of dApps [158], [159]. Figure 5.6 provides a high-level overview of the different O-RAN closed control loops.

Two relevant O-RAN architecture extension efforts that need to be highlighted are currently ongoing in the scope of SNS JU. First, the TERRAMETA project [160], [160], [161], [162] is studying the integration of THz RIS by considering different deployment scenarios. In this case, the integration relies on the definition of new O-RAN entities and interfaces, as can be seen in Figure 5.7.

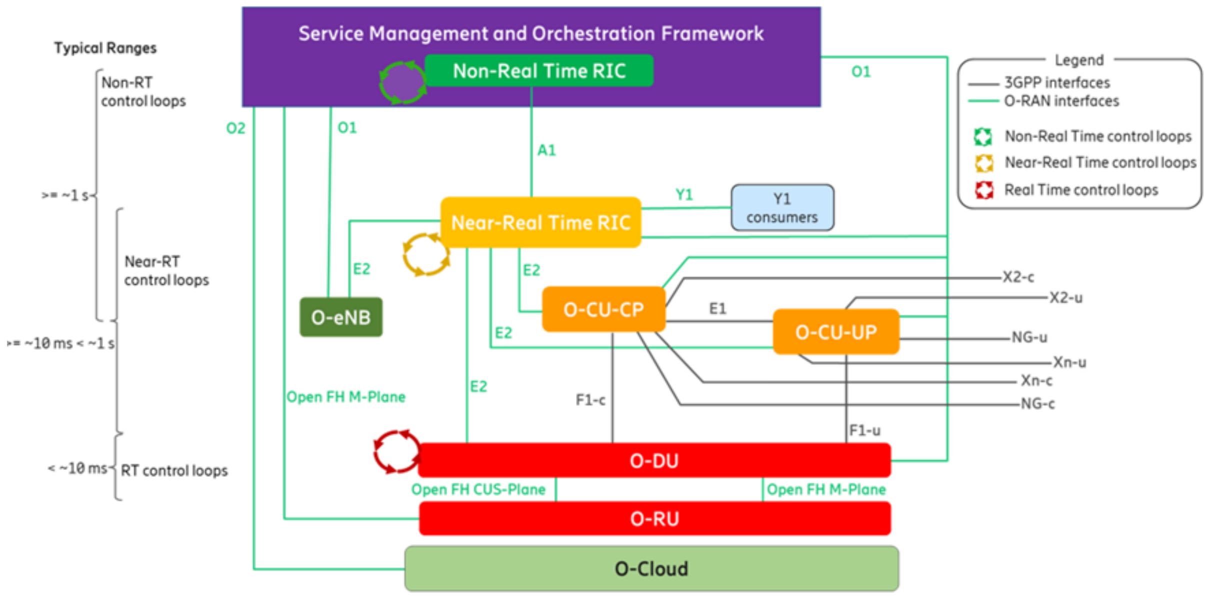

Figure 5.6. High-level overview of the O-RAN closed control loops. [154].

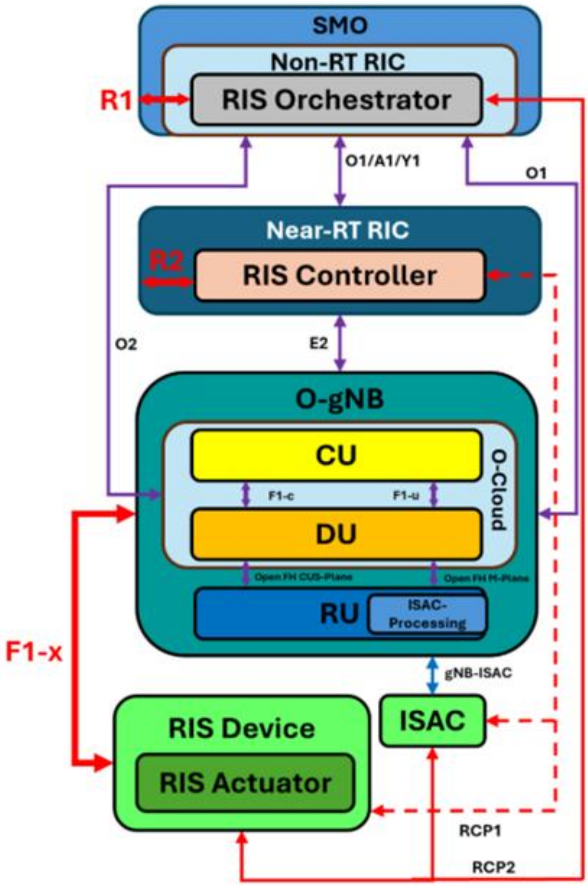

Figure 5.7. New RIS O-RAN interfaces proposed by TERRAMETA [161].

Second, the BeGREEN project [163], [164] also considers the extension of the current O-RAN architecture to integrate RIS, relay devices, edge computing and AI engines for network optimization, as shown in Figure 5.8. Similarly to TERRAMETA, BeGREEN also considers the definition of new interfaces, as well as the extension of the current ones, in their proposed enhanced O-RAN architecture.

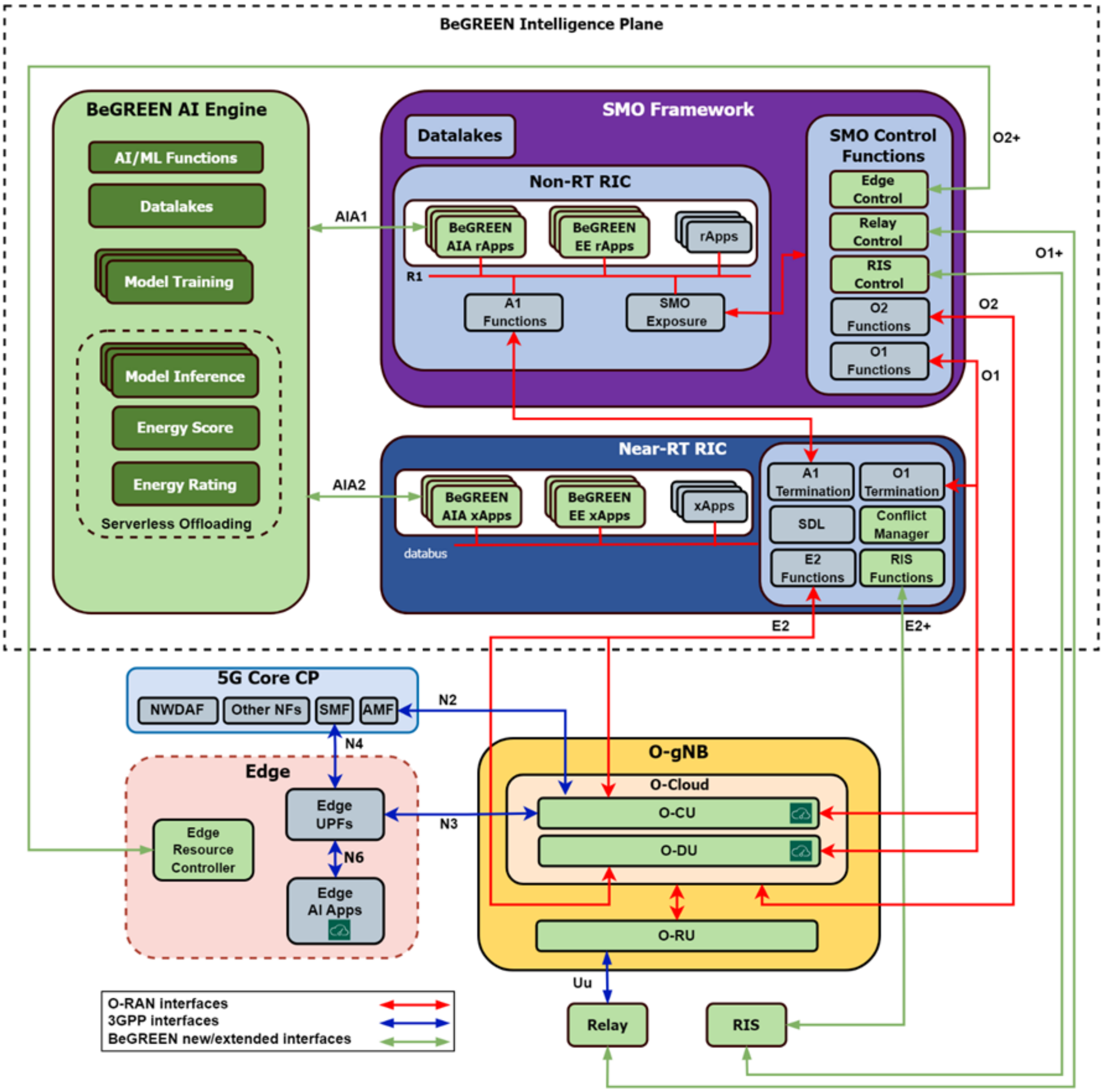

Figure 5.8. Extended O-RAN architecture proposed by BeGREEN [164], [165], [166].

6G-LEADER aims at proposing and evaluating a set of innovative O-RAN architecture enhancements to address the limitations discussed above, while closely following the frequent technical updates published by the O-RAN Alliance [167]. The extensions and enhancements proposed by the SNS JU projects mentioned above will be thoroughly studied and considered when defining **6G-LEADER** 's O-RAN extensions, which aim to further enhance the RAN architecture by embracing innovative concepts such as semantic awareness and sub-10ms control loops. In more detail, the different O-RAN implementations comprising the RAN platforms of **6G-LEADER** (dRAX 5G from ACC, srsRAN Project from SRS) will be enhanced and extended accordingly. Moreover, the integration and interoperability of the extended O-RAN components, between them (e.g., CU from ACC, DU from SRS) and with other relevant **6G-LEADER** innovations (e.g., RU from MB, DU from SRS) is a major priority to the project, as well as promoting the findings for consideration by the relevant bodies (e.g., O-RAN Alliance, 3GPP).

5.2.3 Conflict management in O-RAN

The O-RAN architecture represents a transformative approach to traditional RAN systems, fundamentally reshaping its architecture by fostering a flexible, multi-vendor environment, and eliminating any vendor lock-in. This is achieved through the introduction of open interfaces between disaggregated RAN components, which enable scalable architectural designs and encourage the integration of ML/AI-based custom logic. Custom logic within the O-RAN framework is delivered via rApps, xApps, and dApps, each tailored to perform specific roles in network management. These applications operate autonomously across varying timescales, contributing to the network's decentralized and agile nature. Each app is optimized for distinct tasks and RAN functions, which enhances architectural flexibility and reduces the risk of single points of failure.

Despite these advantages, ensuring optimal RAN performance necessitates careful coordination among these applications. This is particularly crucial when multiple apps operate within the same domain, such as managing shared resources like radio spectrum or computational capacity. Without proper coordination, conflicts may arise, leading to degraded network performance. Such issues can stem from the localized scope of information accessible to each app or from the lack of joint optimization under specific network conditions. Therefore, a strategic approach to application coordination is essential. By aligning their operations and ensuring synergy, the **6G-LEADER** O-RAN architecture will embed a conflict management framework that can maximize performance benefits while maintaining O-RAN openness, flexibility, and efficiency.

5.2.3.1 Overview on O-RAN Conflicts

The lack of inherent awareness among RAN intelligent Apps regarding each other's decisions can lead to potential conflicts between multiple agents providing RAN control across various parts of the architecture. An in-depth analysis of the various types of potential O-RAN conflicts [168] is provided below, and an illustrated overview is provided Figure 5.9, in 1) Intra-Non-RT RIC Conflicts among rApps; 2) Intra-Near-RT RIC Conflicts among xApps; 3) Conflicts among dApps; 4) Inter-RIC Conflicts across same timescale RICs; and 5) Inter-RIC Conflicts across different timescale RICs. Horizontal conflicts are highlighted in yellow. Vertical conflicts in grey.

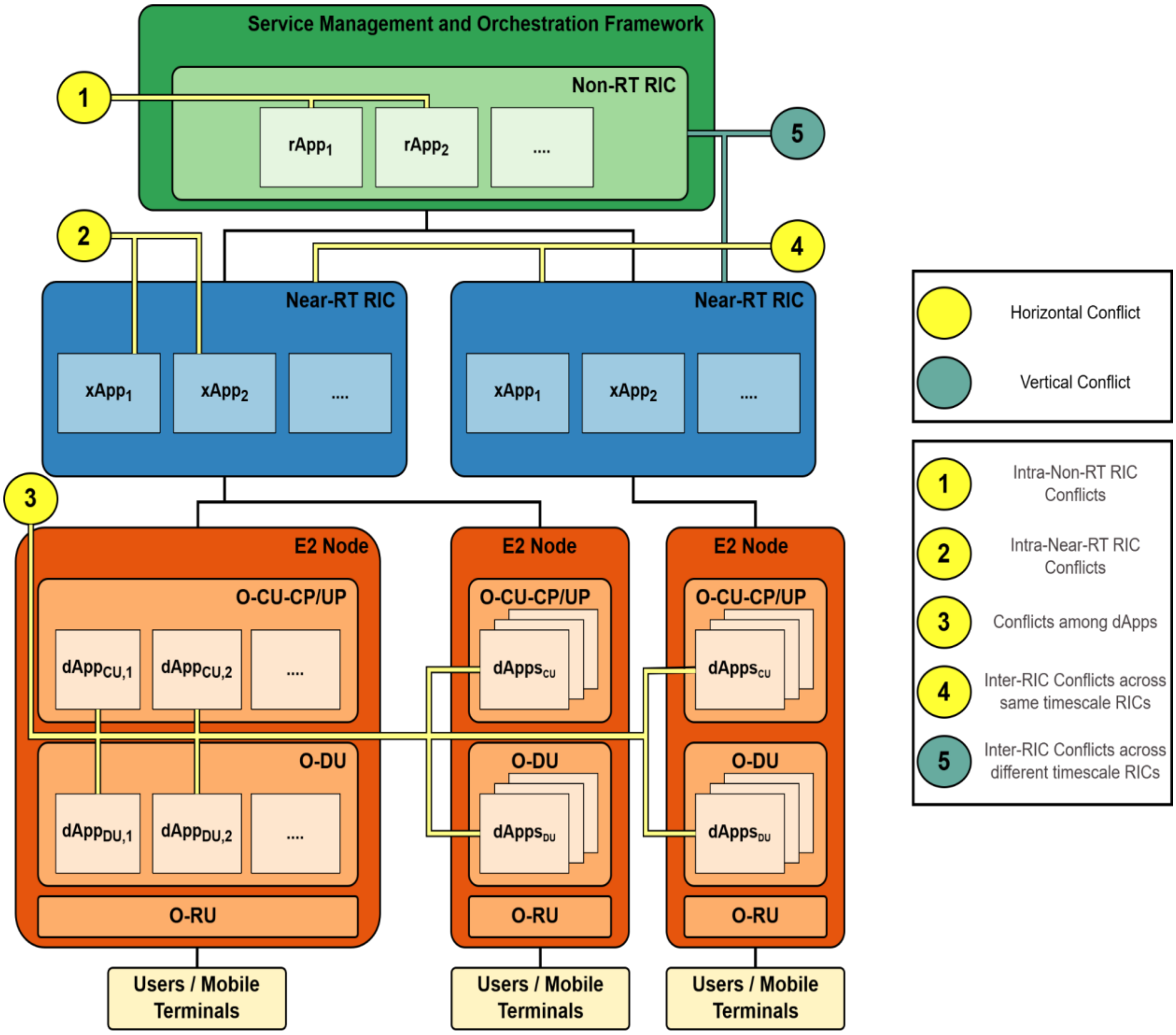

Figure 5.9. Overview of potential O-RAN conflicts.

- *Intra-Non-RT RIC Conflicts among rApps* (Figure 5.9, 1): rApps support and facilitate RAN optimization and operations by providing policy guidance, enrichment information, configuration management and data analytics. Other common examples of rApps include frequency and interference management, RAN sharing and network slicing. However, conflicts can arise when rApps pursue competing objectives—especially when they manage the same resources concurrently. For instance, one rApp may prioritize low latency, while another aims to maximize throughput, leading to potential conflicts in resource allocation. These issues can be further compounded by differences in priority levels. Additionally, rApps developed by different vendors may encounter compatibility challenges due to version mismatches and implementation inconsistencies, such as differing execution timings. These disparities can result in coordination issues and reduced operational efficiency.
- *Intra-Near-RT RIC Conflicts among xApps* (Figure 5.9, 2): Multiple xApps may simultaneously attempt to modify the same control parameters or update different but interdependent parameters that ultimately influence the same network metrics. Such

uncoordinated actions can degrade network performance. For example, one xApp might adjust antenna tilt while another modifies cell offsets, leading to operational inefficiencies. Conflicts are further intensified by the reliance on shared data—differences in how xApps interpret or act upon this data can result in inconsistent outcomes. Additionally, competition for resources such as processing power, memory, and bandwidth can hinder xApp performance, ultimately impacting overall network functionality.

- *Conflicts among dApps* (Figure 5.9, 3): Multiple dApps may simultaneously attempt to modify the same control parameters or adjust different but interdependent parameters that influence critical network metrics in real time. Without proper coordination, these actions can compromise network performance. For instance, one dApp might adjust transmission power while another optimizes beamforming, leading to conflicting adjustments and operational inefficiencies.
  Conflicts among dApps can arise regardless of their deployment location. Intra-O-CU/DU conflicts occur when multiple dApps operate within the same O-CU/DU instance aiming at conflicting objectives. Additionally, inter-O-CU-DU conflicts may emerge when dApps running across different E2 Nodes - whether in the O-CU, O-DU - interact without proper synchronization. Competition for limited resources such as processing power, memory, and low-latency bandwidth can degrade the dApp performance and complicate coordination, affecting overall network stability and efficiency. Importantly, conflicts among dApps and their classifications have not been formally standardized or addressed in any research reports by the O-RAN Alliance. As such, these considerations highlight an area where further investigation and standardization efforts may be beneficial.
- *Inter-RIC Conflicts across same timescale RICs* (Figure 5.9, 4): In O-RAN architecture, Inter-RIC conflicts can arise between multiple RICs operating at the same timescale and independently managing overlapping or adjacent RAN segments, leading to contradictory control decisions. Such conflicts can manifest through inconsistent policy enforcement, conflicting parameter adjustments, or resource allocation inefficiencies. For instance, Near-RT RICs might trigger opposing handover decisions, while RT RICs could create interference by making unsynchronized adjustments to radio parameters. Non-RT RICs may introduce conflicts through divergent long-term policies or inconsistent machine learning model updates. These challenges are compounded by variations in data interpretation, resource competition, and timing mismatches.
- *Inter-RIC Conflicts across different timescale RICs* (e.g., between Non-RT and Near-RT RICs) (Figure 5.9, 5): Additional conflict risks emerge due to the multi-timescale control loops in O-RAN. Although apps operate across different timescales, they often rely on the same data or resources, creating potential for conflict. For example, an xApp may prioritize URLLC users, while an rApp focuses on optimizing cell load for long-term throughput, resulting in contradictory network commands that negatively affect sensitive users. Inconsistent interpretations of shared data can lead to misaligned actions, while differences in interface standards or protocols between these applications can further cause operational conflicts.

The conflicts between equivalent components can be grouped and referred to as horizontal conflicts (highlighted in yellow in Figure 5.9), or vertical conflicts that are identify the conflicts between components on different control levels of the architecture (highlighted in grey in Figure 5.9).

The O-RAN Alliance has also introduced another conflict classification – direct, indirect, and implicit – based on how the decisions made by xApps can interfere with each other [168]. This classification can be easily extended to all the types of conflicts presented above and illustrated in Figure 5.9.

- *Direct Conflicts:* These occur when two xApps make contradictory decisions that affect the same set of configuration parameters, leading to one decision overriding the other. An example is when xApp1 assigns a user to a specific cell, and xApp2 assigns the same user to a different cell. If these conflicts go undetected, xApp1 could draw wrong conclusions about the impact of its actions, leading to network inefficiency. Direct conflicts can happen whenever two xApps are making decisions that directly affect the same resource or entity without coordination.
- *Indirect Conflicts:* These occur when xApps make decisions that influence overlapping or related areas of the RAN operation, but not necessarily the same parameters. These decisions may lead to uncoordinated and fluctuating outcomes in the network. For example, if xApp1 adjusts the electrical tilt of an antenna, while xApp2 modifies the Cell Individual Offset (CIO), these changes might cause inconsistent handover boundaries, as the actions of one xApp interfere with the other, resulting in suboptimal performance. Indirect conflicts arise when the decisions made by different xApps affect interconnected system components or parameters, even if they aren't directly changing the same settings.
- *Implicit Conflicts:* Implicit conflicts occur when xApps optimize the RAN for separate, often competing, objectives, which might lead to contradictory outcomes even though the xApps are not directly modifying the same parameters. For instance, if xApp1 focuses on maximizing the QoS for a group of users, while xApp2 aims to minimize the number of handovers between neighbouring cells, these conflicting goals may lead to a situation where the decisions of one xApp negatively affect the objectives of the other, disrupting network performance. Implicit conflicts arise when different xApps are working towards distinct, sometimes conflicting, goals that influence the overall network operation, even though they may not be directly interacting with the same parameters.

5.2.3.2 Challenges in Conflict Detection and Management

Conflict Detection and Management (CDM) within the O-RAN architecture is complex due to its open, disaggregated, and multi-vendor nature. The main challenges associated with the design and implementation of the CDM framework include:

- *Multi-vendor interoperability:* O-RAN promotes a multi-vendor ecosystem. Hence, the implementation of functions and assumptions within the developed applications (i.e. xApps/rApps) and their configuration could be different, e.g. for the policies defined within Non-RT-RIC and related enforcements within Near-RT RIC. Also, the applications could be trained with different data sets and therefore, they result in different actions, i.e. predictions for the given scenario and subsequent control actions. Due to these differences between interpretation of implementations for the applications as well as the potential inconsistencies and contradicting behaviours, it is very difficult to formulate the problems for the design of effective CDM mechanisms.

- *Specific requirements per deployment scenarios and use cases:* The requirements for the future networks and use cases are very diverse and in practice the trade-off for the competing KPIs is typically defined for specific deployment scenario. From a network design point of view, even defining the optimal state of operation for the given use case and deployment scenario based on the available set of competing KPIs might be a difficult task. The use case and deployment-specific requirements will be further customer-specific in the future, particularly within the Private Networks domain. These factors make the design of effective CDM solutions highly customer-specific and therefore adds further complexity for the O-RAN community to design universal CDM frameworks (particularly for the developers to develop one-size-fits-all solutions).
- *Definition and awareness of network state: for effective operation of CDM:* The CDM needs to have a frequent and consistent view on network state, which is very difficult to attain in practice. The control loop for different applications within the O-RAN architecture is different, e.g. xApps operates in near real-time (10ms–1s), while rApps operate in non-real-time (>1s) [49], [154]. Hence, specifying the coordination mechanisms to precisely interpret and report the overall state of the network and avoiding reporting of outdated state is very difficult. The situation is further exacerbated in large-scale deployments with many applications running since the number of possible interactions within the network will grow rapidly. In such cases, new challenges might be introduced that do not allow the system to operate efficiently. For example, conducting coordination and optimisations for CDM to come up with a set of decisions and actions that consider collective interest(s) of different competing factors can be computationally intensive. Therefore, running such coordination(s) can be potentially counterproductive due to the negative impact on network's responsiveness to changes, e.g. competing factors can be within contradicting objectives between time-sensitive decisions and long-term policies that need to be addressed while network conditions are dynamically changing. The advantages and potential gains of O-RAN Apps are mainly discussed and validated when deployed as standalone solutions while the required alignments between the different Apps (e.g. synchronisation between different Apps with different time scales that could impact each other's decisions and resulting actions) have not been explored widely and are still in their infancy, e.g. an early design of a CDM framework was proposed in [169], but the design was limited for managing conflicts among xApps within Near-RT RIC, and the validations were conducted within an emulated O-RAN network.
- *Security implications and trust:* Within the xApp(s)/rApp(s) and CDM, malicious or poorly designed and tested solutions could result in harmful decisions and control actions. Hence, while development of applications and CDM aims to improve and optimise the performance and operation of the network, it could result in opposite and destructive outcomes, e.g. cyber-attack to the CDM unit could result in making adjustments based on attacker's desire and against original intended factors, particularly the CDM entity could

disclose information about the desirable configuration for the overall operation of the system and priorities and preferences. Hence, this tightens the requirements for the establishment of security and trust specifications, guidelines and practices to ensure trustworthy behaviour and interactions for CDM. Also, it is one of the bottlenecks in practical utilisation of xApp(s)/rApp(s) and CDM applications within real-life deployments.

- *Complexities for performance validations:* The practical gain of O-RAN applications and CDM can only be demonstrated when they are deployed in various commercially neutral platforms (e.g. existing Open Testing and Integration Centres, OTICs [170]) that allow to test and validate interoperability of various products for different vendors and to present the gains that such plug-and-play solutions could offer. However, the reality of O-RAN for the future carrier-grade networks is not well established and has not reached maturity due to several reasons, e.g. the lack of clarity about the common set of technical requirements to confirm and validate product readiness for real-life network deployments. This will further hinder the progress required for developing the applications and CDM solutions that could yield practical gains in multi-vendor networks. This requires conducting testing, validation, and certification programmes for applications and CDM solutions like the existing programmes for collaborative testing of the O-RAN components, i.e. O-RU, O-DU, O-CU, and RIC.

In summary, the above-mentioned challenges need to be addressed for the development of CDM solutions for the O-RAN to offer practical gains in real-life networks. O-RAN Alliance provides specifications for interfaces and guidelines for interoperability testing. Recently, O-RAN initiated CDM focused standardisation activities and published the first version of technical report [171] about Conflict Mitigation functions. However, this document only covers the background knowledge in this domain (e.g. type of conflicts) and addresses a few specific issues within conflict detection, resolution, and avoidance between the xApps within the Near-RT RIC. Still, there is not any enforced guideline, standard, or set of practices for the design of CDM that considers both Near-RT RIC and Non-RT RIC requirements and interactions. The ongoing R&D within O-RAN ecosystem enables acceleration of the developments needed for CDM solutions by identifying functionalities needed in that space which can gradually shape the standardisation activities, e.g. in [172], authors conducted the study of how different use cases can work harmoniously within O-RAN architecture, presented a framework designed to handle the xApps-based network management optimisations, and shared implementation roadmap for the development of such functionalities.

5.2.3.3 O-RAN Conflict Detection and Management Initiatives

The following subsections delve into the analysis of existing conflict detection and mitigation, which have been the focus of investigation in both European research projects and research initiatives.

5.2.3.3.1 Existing frameworks

Recently, several frameworks have surfaced to address the challenges posed by conflict detection and resolution in an O-RAN environment. However, these advancements have generally been narrowed in coverage and fall short of the intelligence required for efficient and adaptable conflict management in disaggregated systems such as O-RAN.

The main methodology used relies on a steady partitioning of control scopes, where several xApps were assigned to specific sets of RAN parameters. In this way, limitations and constraints started to become apparent as RAN intelligence evolved toward multi-agent systems with real-time decision-making.

These limitations have been acknowledged by the O-RAN Alliance, which emphasized in its technical specifications the necessity for conflict awareness in the Near-RT RIC [171]. In particular, this relates to lifecycle control management, policy enforcement, and coordination between xApps, and dApps operating on the same resources. As a result, some research began investigating intent-based conflict tagging [173], where control messages were enhanced with additional metadata to describe the action being performed, the RAN components it targets, and the expected duration or type of impact. This was efficient in adding some traceability. However, the problem was that the frameworks might not realize and react to these tags intelligently in real-time. As a result, conflict detection was in most cases rule-based, depending on static, manual policies to filter or override control commands based on predefined conditions.

Indirect and implicit conflicts represent subtle yet significant challenges, often arising from the interaction of independent control loops that influence shared KPIs. Only a limited number of early frameworks attempted to address these issues through correlation-based diagnostic methods. These approaches sought to associate observed performance degradations with preceding control actions by analysing historical logs to infer potential causal relationships. For instance, a conflict detection mechanism is introduced within the Near-RT RIC [169], relying on rule-based analysis of message flows to spot conflicts. While the concept showed promise, these mechanisms were largely reactive, offering limited support for real-time intervention or proactive conflict resolution.

Another significant limitation of these frameworks lies in their lack of memory and contextual awareness. Most operated without maintaining a persistent history of previous conflicts or the strategies used to resolve them. As a result, they were unable to identify recurring patterns or adapt their conflict resolution mechanisms over time. Recent attempts proposed the use of graph neural networks to reconstruct and learn from conflict structures based on past xApp behaviour, yet they remained constrained to conflict inference rather than real-time mitigation [174]. Similarly, the proposed framework [175] relied on pre-deployment profiling in a sandbox environment to catch possible xApp conflicts ahead of deployment. But it still lacked the ability to coordinate responses at runtime and didn’t include a dynamic feedback mechanism during execution.

Additionally, many implementations lacked a dedicated runtime component, such as a centralized conflict manager that could dynamically coordinate conflict detection and resolution across multiple xApps in real time [176].

Even with the progress made in areas like intent tagging, and AI-based post-analysis, most of the existing frameworks are still quite fragmented. They tend to be reactive and often tied to just one part of the xApp lifecycle. What's usually missing is the ability to understand the context in real time, remember past conflicts, or coordinate actions, capabilities that are essential in dynamic, low latency RAN environments involving multiple domains. Additionally, they still fall short when it comes to handling implicit and indirect conflicts, especially in live scenarios where independently developed xApps are running at the same time.

To address the limitations of prior approaches, **6G-LEADER** introduces a novel Conflict Manager as a core component of its architecture. This module is designed to manage semantically enriched xApps operating across diverse control and optimization loops within the near-RT RIC. Unlike earlier frameworks that relied on static rule enforcement or post-action diagnostics, the **6G-LEADER** Conflict Manager proactively evaluates incoming control messages and determines whether specific network reconfigurations should be permitted or blocked. Importantly, the framework is also equipped to manage inter-RIC conflicts, a capability that is notably absent in most existing solutions. This is achieved through a coordinated information exchange mechanism between multiple near-RT RICs, orchestrated via the non-RT RIC.

Additionally, with the expected deployment of dApps in future RAN environments, the framework anticipates tighter latency constraints and more critical conflict resolution timelines. To meet these demands, **6G-LEADER** incorporates pre-action conflict resolution capabilities, ensuring that conflicts can be detected and addressed before control actions are executed. This marks a significant shift toward proactive, real-time conflict avoidance in next-generation O-RAN systems.

5.2.3.3.2 Existing EU projects handling Conflict Mitigation

Several SNS-JU European projects including ERGE, ACROSS, 6G-INTENSE and ETHER have been working on conflict detection and resolution:

- *VERGE* [177]: This project works on resolving the challenges associated with the widespread adoption of multiple, independent AI/ML solutions across the VERGE system. Specifically, the design of multi--level, multi-agent mechanisms for the coherent integration between some of the AI components are investigated. Such mechanisms detect and resolve potential conflicts between the independent AI/ML--based decisions, ensuring that the system behaviour is stable and efficient. **6G-LEADER** will adapt conflict mitigation solutions between the AI model decisions from the VERGE project and enhance the conflict detection/mitigation solution within the O-RAN conflict management framework instead of concentrating on the detection and resolution only with no O-RAN conflict management framework.

- *ACROSS* [178]: The project provides a detailed analysis of the requirements for an automation platform and software development kits that will enable CSPs and app developers to be successful, as well highlighting the need for conflict detection and mitigation between independent automation processes. Particularly, ACROSS focuses on conflict mitigation between rApps at the SMO as defined by O-RAN Alliance. **6G-LEADER** will follow the conflict detection and mitigation work from ACROSS on rApps closely. Additionally, our solution will provide a comprehensive framework covering not only rApps but also xApps at near-RT RIC and demonstrate it in a PoC.

- *6G-INTENSE* [179]: 6G-INTENSE works on conflict detection and mitigation between multiple intents, particularly focusing on the faults and workload variations to test the system's adaptability, intra and inter–domain coordination and conflict resolution. The project aims to autonomously mitigate over 90% of the conflicts arising from multi-domain deployment. **6G-LEADER** will closely monitor the findings from 6G-INTENSE, and additionally, the project will develop a conflict management framework at near-RT RIC, focusing on the conflict detection and mitigation between multiple xApps, dynamically changing RAN parameters. Extensive list of KPIs will be defined to measure performance of different parts of the conflict mitigations frameworks (e.g., detection accuracy and reporting success rate, conflict resolutions effect on system performance improvements and conflict avoidance success rate).

- *ETHER* [180]: This project designs a Network Intelligent Orchestrator (NIO) that is responsible for supervising the Life Cycle Management (LCM) of the Network Intelligent Services (NIS) by efficiently harmonising the Network Intelligent Functions (NIFs) that constitute each of them. So, the NIO works on several main purposes when it comes to the efficient coexistence of AI models across network domains and planes, including conflicts avoidance, such as the ones generated by the existence of different NI algorithms that aim to configure the same network functions or resources that run at various timescales or based on diverse input. In **6G-LEADER**, the O-RAN Alliance conflict management framework will be advanced and expanded to include conflict detection and mitigation, alongside avoidance strategies.

- *BEGREEN* [165]: In the BeGREEN project, conflict mitigation extends over several RAN areas such as SMO, Non-Real-Time RIC, Near Real-Time RIC, and interfaces. To support general conflict management in the RIC, the project targets some modifications for the dRAX framework. For example, the non-RT RIC needs to define an A1 policy manager, extending the Near-RT RIC to support several entities, which can be divided into three specific areas: Subscription manager, conflict manager and conflict avoidance handler. The solution is based on collaboration, where in this collaborative approach, if two xApps attempt to use the same resource, they step back and resolve the conflict using information from the dRAX databus based on information from the policy. **6G-LEADER** will closely monitor the findings from BeGREEN project, and additionally, will advance and expand the framework to include mitigation in the use case where conflict avoidance is not possible between multiple xApps due to dynamically changing RAN parameters. Moreover, an extensive list of KPIs will be defined to measure performance of different parts of the conflict mitigations frameworks.

In all the above projects, there is no specific conflict management framework for near-RT RIC focusing on RAN actions conflicts in various timescales. In **6G-LEADER**, all above projects and conflict management work in O-RAN Alliance will be closely monitored and in addition, a comprehensive conflict management framework will be worked on enhancing the current O-RAN Alliance architecture. Additionally, developed conflict management solutions will be showcased in a PoC and KPIs will be collected and evaluated.

6 6G-LEADER High-Level RAN Architecture

The high-level RAN architecture presented in this deliverable represents the preliminary design phase of **6G-LEADER** architectural framework. Building upon the modular principles of the O-RAN standard, it integrates a set of technological innovations that aim to address the extreme requirements foreseen for future 6G networks, including ultra-low latency, high energy efficiency, and AI-native network intelligence. At this stage, the architecture should be regarded as an initial blueprint to provide a structured foundation that guides the ongoing research activities within the project while saving room for progressive refinement as the work advances. Ath the end, the final architecture would be applicable to only O-RAN base architectures but to wider ones.

The preliminary version of **6G-LEADER** architecture (Figure 6.1) reflects the insights and objectives defined so far in WP2 and in particular Task 2.4, where the architecture is being actively evaluated and shaped considering emerging findings and technical progress. It brings together the main functional elements of the RAN and introduces new capabilities such as semantic-aware communication, real-time control loops, and advanced antenna technologies. Furthermore, the present architecture is conceived as part of an iterative design process. Future deliverables will revisit and extend this initial proposal to incorporate feedback from PoCs implementation, performance evaluations, and alignment with standardisation efforts.

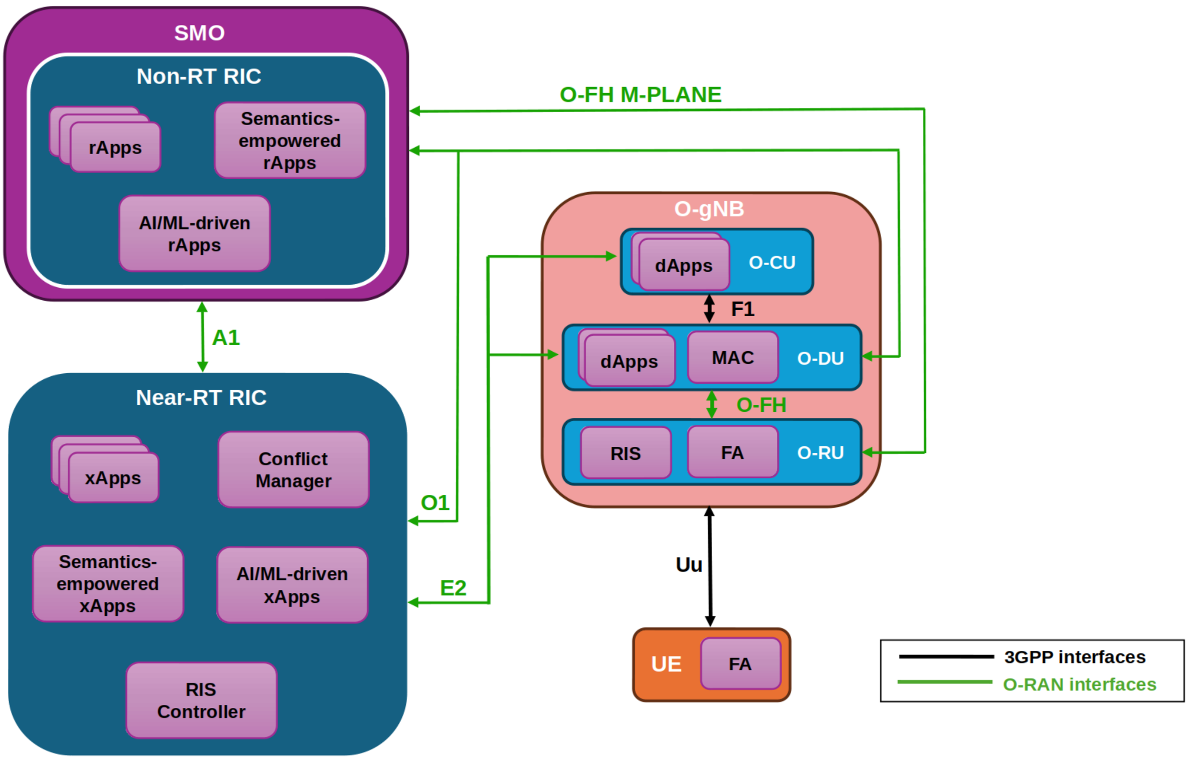

Figure 6.1. High-level 6G-LEADER architecture.

6.1 How the requirements describe the RAN components

The design of the integral RAN architecture in **6G-LEADER** is grounded in a comprehensive set of technical, functional, and non-functional requirements that reflect the expected demands of 6G services and infrastructure. These requirements have been methodically mapped to architectural principles, component specifications, and control mechanisms to ensure that the resulting RAN system is well aligned with long-term technological evolution and feasible for real-world deployment.

To manage the increasing complexity and scale of 6G networks, the RAN architecture must embed native support for AI/ML, enabling predictive, optimization, adaptive, and data-driven control mechanisms across all operational layers. This requirement will be fulfilled through the integration of AI/ML models across the Non-RT, Near-RT, and RT domains. The architecture includes dedicated interfaces and MLOps workflows to support model deployment, validation, and continuous adaptation based on real-time network observations, ensuring that control decisions are still accurate and responsive under evolving conditions.

Future 6G RANs must achieve significant reductions in power consumption and EMF emissions, aligning with sustainability KPIs. This requirement will be fulfilled through the deployment of highly RF components, such as FAs and RISs, that enable context-aware beamforming, and spatial resource management and orchestration, as well as environmentally adaptive transmission strategies.

Applications such as XR, industrial control, and autonomous systems demand sub-10ms responsiveness and deterministic behaviour. This requirement will be fulfilled through a third, real-time closed-loop control layer implemented by dApps at the O-CU/O-DU levels, enabling time-critical RAN functions to be executed with minimal latency.

The RAN must support massive device densities, highly heterogeneous traffic profiles, and distributed computing workloads without degradation in QoS. This requirement will be fulfilled through leveraging AI-enhanced multiple access schemes, such as NOMA and RSMA, that facilitate dynamic, fine-grained resource allocation and interference management.

Within the O-RAN, maintaining interoperability among diverse xApps/dApps and preventing operational conflicts is critical for stable RAN performance. This requirement will be fulfilled through the integration of a dedicated Conflict Manager within the Near-RT RIC, providing coordinated decision-making and safeguarding the overall consistency of network operations.

To reduce redundant signalling, enhance task-context alignment, and enable meaningful data prioritization, the RAN must incorporate semantic intelligence. This requirement will be fulfilled through the embedding of Semantic intelligence within the control and user planes through advanced xApps and dApps, enabling goal-oriented communication and prioritizing information based on its task relevance. In doing so, semantic-aware control reduces signalling overhead, aligns network actions with application intent, and contributes to the broader vision of AI-native, context-aware 6G networks.

6.2 Integral RAN architecture proposal

The integral RAN architecture proposed by **6G-LEADER** enhances and further develops the O-RAN architectural framework to address the stringent performance and flexibility demands expected in future 6G networks. A key architectural element is the deployment of closed-loop control mechanisms operating across three distinct time domains: Non-RT, Near-RT, and RT. Each control loop is supported by specialized AI/ML components and semantic processing techniques, working collaboratively to optimize the management of communication and computation resources within the RAN.

At the non-RT layer, long-term learning and optimization tasks are carried out by rApps operating within the SMO domain. These tasks facilitate the training and continuous improvement of AI models utilized by downstream controllers, enabling the enforcement of global policies and the provision of resource management guidance based on historical data patterns and operator-defined objectives. The non-RT RIC acts as a repository and training environment for AI/ML models, which are subsequently distributed to downstream controllers for inference, enabling the system to adapt to evolving network contexts while remaining aligned with long-term service-level goals.

The Near-RT control loop is implemented within the RIC, where xApps utilize AI/ML techniques and semantic awareness to dynamically optimize radio resource management at sub-second timescales. These xApps leverage predictive models and real-time key performance metrics to manage complex functionalities such as user scheduling, link adaptation, and beamforming. These enhancements aim to improve energy efficiency and control EMF exposure by leveraging mechanisms such as reconfigurable RF components, including FAs and RIS, as well as power control strategies and hybrid beamforming techniques. Semantic information and deep reinforcement learning are used to inform real-time decisions on FA port activation, RIS phase shifting, and adaptive power/rate allocation, particularly in the context of non-orthogonal multiple access schemes.

To enhance computational resource efficiency within the RAN, the proposed architecture adopts a semantically informed task allocation approach. In this context, computational tasks are dynamically assigned to the most appropriate execution nodes (e.g., edge UEs, IoT devices, or distributed edge clusters) based on a combination of task descriptors, QoS constraints, and the real-time system state. This facilitates the realization of the "Wireless for AI" paradigm, which integrates AirComp techniques to support federated learning and distributed inference tasks directly within the RAN. By combining communication and computation in a unified, semantically optimized control framework, the architecture ensures scalable, low-latency, and energy-efficient service delivery.

Beyond near-RT operations, **6G-LEADER** incorporates a third, RT control loop that operates at the level of the O-DU and O-RU, supported by containerised dApps, designed to deliver ultra-low-latency inference and control capabilities. The dApps facilitate the execution of time-critical functionalities such as physical layer scheduling and low-level beam adaptation, with reaction times in the sub-10ms range.

To ensure coherent and conflict-free operation across all control loops, **6G-LEADER** introduces a dedicated Conflict Manager within the near-RT RIC. This module actively oversees control decisions generated by multiple xApps and dApps, identifies potential direct, indirect, or implicit conflicts, and applies appropriate resolution mechanisms either proactively or reactively. The Conflict Manager leverages a historical knowledge base and semantic parameter mappings to enable context-aware decision making and preserve RAN operational consistency.

As with the broader architectural design, this integrated control framework represents a preliminary proposal that will evolve through iterative refinement. Subsequent deliverables will incorporate experimental feedback, PoCs results, and standardization inputs, progressively converging toward a fully validated 6G RAN architecture capable of seamless, intelligent, and sustainable network operation.

7 Conclusions

This deliverable ***D2.1 Use case analysis, KPIs and requirements to RAN architecture design***, meets the objectives form *Task 2.1: Technology radar and baseline technologies identification* and *Task 2.2: 6G PoCs, sustainability and requirements analysis*. T2.1 defined a description of the main technical challenges and the technologies that can support the objectives of **6G-LEADER** and how they can be implemented in the PoCs defined. It monitored relevant trends and catalogued baseline technologies that impact **6G-LEADER**, identifying wireless communication and signal-processing enablers for integration into the O-RAN–based architecture. In addition, a comprehensive state-of-the-art was presented spanning across PHY/ML, semantics, reconfigurable RF (FAs/RIS), multiple access (incl. AirComp), and O-RAN extensions (xApps/rApps/dApps, conflict management) with interface implications, providing inputs to future architecture design. Meanwhile, T2.2 defined a top-down method to move from high-level intent to measurable results, by mapping that method to concrete requirements and KPIs, and by presenting an initial architecture. It starts from societal drivers and SDG alignment, identifies the innovation areas where **6G-LEADER** must act, turns those areas into project objectives and high-level use cases, quantifies success through KPIs and E2E requirements, and connects all of this to the first version of the RAN architecture and to a set of PoCs. In doing so, it establishes the traceability chain that the project will use in design, integration, and validation.

Chapter 2 sets the methodology. It formalises the flow from SDGs to innovation pillars, objectives, use cases, KPIs, and E2E requirements. Each step produces items that can be verified later: SDG alignment statements, pillar scope notes, objective statements, use-case briefs, KPI targets, and requirement lists. The chapter also anchors the KPI families to the SNS white paper so that definitions and targets are comparable with the wider community. Hence, the method can absorb changes from later research and validation without losing traceability. As a result, the project has a consistent way to justify design choices and a practical basis for planning tests. Building on that, Chapter 3 performs the mapping. It links SDGs to the project's seven innovation pillars and derives the corresponding objectives. It then binds each objective to one or more technical KPIs and associates those KPIs with the UCGs that will exercise them. The mapping tables are not only descriptive; they define the acceptance conditions for later phases. Each UCG now carries a clear link to the pillars it touches, the objectives it advances, and the KPIs it must demonstrate via the defined PoCs. This reduces ambiguity when specifying scenarios, traffic profiles, and measurement points, and it prepares a clean handover to integration and validation activities.

Chapter 4 identifies the AI-driven advanced communication techniques that will move the KPIs. AirComp is positioned to lower aggregation latency and radio overhead for distributed learning and control. Semantic communications targets reductions in non-useful traffic and improvements in timeliness metrics (e.g., AoI). AI/ML-aided multiple access and predictive scheduling address spectral and energy efficiency. Near-RT and non-RT RIC logic with conflict management provides the control hooks to turn per-link or per-function gains into system-level effects. The chapter clarifies where each technique is expected to help, how it maps to KPIs, and what test hooks are needed for later measurement. In parallel, Chapter 5 treats reconfigurable components as primary elements of the system. Fluid antennas and RIS, together with FR1/FR3 coexistence, are scoped for their expected gains in spectral efficiency, EMF exposure, and energy use. Just as important, the chapter defines how these components will be steered and observed: configuration interfaces, calibration steps, and closed-loop control so that the RIC can orchestrate them and the test infrastructure can measure their impact. This turns hardware features into controllable resources inside the network rather than isolated lab assets.

Chapter 6 ties all the elements together in the initial **6G-LEADER** RAN architecture. It places RU/DU/CU splits, defines data and control planes for AI-native operation, and locates near-RT and non-RT RIC functions that will host optimisation logic and policy. It also outlines telemetry and data pipelines for training and inference and shows how reconfigurable RF elements connect into end-to-end control loops. The design points back to the requirements and KPIs established earlier, so the contribution of each functional block to project targets is explicit. The architecture is detailed enough to host the planned PoCs and to integrate enabling components from the research work packages, while leaving room for iteration based on measurements.

The relationships with other work packages follow directly from this structure. WP2 supplies the upstream contract: SDG alignment, objectives, KPI targets, and requirement sets. WP3 uses these to guide AI/ML-enhanced PHY/MAC and over-the-air computing; WP4 develops this sustainable 6G RAN new technologies for Fluid Antennas and RIS, while WP5 applies them to goal oriented semantic empowered communication for operation efficiency and sustainability use cases. WP6 consumes the artefacts to consolidate the system view, with a coherent extension for AI/ML methods and semantic extensions to the proposed 6G architecture. WP7 plans scenarios, instrumentation, and success criteria using the PoC-to-KPI links defined here, then feeds results back so WP2 can update the architecture without breaking traceability. This creates a closed loop where design, integration, and validation stay aligned and controlled via WP7.

Looking ahead, WP2 will extend this deliverable through Task 2.4. The architecture presented in Chapter 6 will be expanded into a full specification set, including interface definitions, deployment blueprints across the target testbeds, and common KPI model so results are comparable across PoCs. Task 2.4 will refine security and privacy aspects of the control loops; describe how semantic processing and AirComp pipelines are provisioned, monitored, and benchmarked; detail integration of FR3-capable RU/DU nodes and RIS/FA control into the RIC; and define how conflict-management policies interact with energy- and traffic-optimising xApps. In parallel, WP2 will maintain the requirements and KPIs and update the validation playbook used by WP6 and WP7. By the end of Task 2.4, the current architecture baseline will be an implementation-ready blueprint with clear acceptance criteria, enabling faster integration and confident evaluation.

Taken together, this deliverable D2.1 provides the method, the mapping, the mechanisms, the hardware levers, and the system view needed to execute the project. The links from objectives to KPIs to PoCs make progress measurable. The architectural baseline makes integration feasible. The handover to other work packages is clear, and the next steps in WP2 are defined. This gives the project a stable foundation and a practical path to demonstrate results on real platforms.

8 REFERENCES

[1] UN, 'The UN Sustainable Development Goals – UN Environment Management Group'. Accessed: May 24, 2024. [Online]. Available: https://unemg.org/our-work/supporting-the-sdgs/the-un-sustainable-development-goals/

[2] 6G SNS, 'SNS JOURNAL/2025', HORIZON-JU-SNS-2024-STREAM-CSA-01, May 2025. Accessed: Sep. 09, 2025. [Online]. Available: https://smart-networks.europa.eu/wp-content/uploads/2025/05/sns-journal-2025-web-1.pdf

[3] I. Patsouras *et al.*, '6G KVIs – SNS Projects Initial Survey Results 2025', Apr. 2025, doi: 10.5281/ZENODO.15220945.

[4] I. Mesogiti *et al.*, '6G KPIs - Definitions and Target Values', Mar. 2025, doi: 10.5281/ZENODO.14621168.

[5] FIDAL, 'D2.1 Requirements, Architecture and Methodologies.', Jun. 2023. Accessed: Aug. 25, 2025. [Online]. Available: https://fidal-he.eu/sites/default/files/pd/file/2023-07/D2.1_FIDAL.pdf

[6] Hexa-X, 'Main - Hexa-X'. Accessed: Apr. 19, 2024. [Online]. Available: https://hexa-x.eu/

[7] S. N. S. TrialsNet, 'TrialsNet', TrialsNet. Accessed: May 31, 2024. [Online]. Available: https://trialsnet.eu/

[8] FIDAL, 'Key Performance Indicators (KPI) Anex 1 – 5G KPIs.', Oct. 2024.

[9] FIDAL, 'Key Value Indicators (KVI)', Oct. 2024.

[10] HEXA-X-II, 'D1.3 Environmental and social view on 6G', Mar. 2023. [Online]. Available: https://hexa-x-ii.eu/wp-content/uploads/2024/03/Hexa-X-II_D1.3_v1.00_GA_approved.pdf

[11] HEXA-X-II, 'D5.5 Final design of enabling technologies for 6G devices and infrastructure', Mar. 2025. [Online]. Available: https://hexa-x-ii.eu/wp-content/uploads/2025/04/Hexa-X-II_D5.5_v1.0.pdf

[12] Hexa-X-II, 'HEXA-X-II - Deliverable D2.2: Foundation of overall 6G system design and preliminary evaluation results', Dec. 2023. [Online]. Available: https://hexa-x-ii.eu/wp-content/uploads/2024/01/Hexa-X-II_D2.2_FINAL.pdf

[13] G. Scivoletto *et al.*, 'Deliverable D2.3 - Final design of Platforms and Networks solutions', Dec. 2024, doi: 10.5281/ZENODO.14512906.

[14] O. M. Gardin *et al.*, 'Use Cases definition for Infrastructure, Transportation and Security & Safety (ITSS)domai', May 2023, doi: 10.5281/ZENODO.7944484.

[15] D. Tsakanika *et al.*, 'Use Cases definition for eHealth and Emergency (eHE) domain', May 2023, doi: 10.5281/ZENODO.7944691.

[16] TARGET-X, 'TARGET-X | Project | HORIZON - Trial PlAtform foR 5G EvoluTion – Cross-Industry On Large Scale', CORDIS | European Commission. Accessed: Aug. 25, 2025. [Online]. Available: https://cordis.europa.eu/project/id/101096614

[17] TARGET-X, 'D1.1 Forward looking use cases, their requirements and KPIs/KVIs (opens in new window)', Dec. 2023. [Online]. Available: https://ec.europa.eu/research/participants/documents/downloadPublic?documentIds=080166e5f7113337&appId=PPGMS

[18] 'Home', OpenTAP. Accessed: Sep. 09, 2025. [Online]. Available: https://opentap.io

[19] 'Getting Started | OpenTAP'. Accessed: Sep. 09, 2025. [Online]. Available: https://doc.opentap.io/

[20] IMAGINE-B5G, 'D1.1 Use cases, Ecosystem Analysis and KVIs/KPIs for OC1 and OC2', Dec. 2024. [Online]. Available: https://drive.google.com/file/d/1J-jbfr7-SJgbcYkQKIu2ZgqwaaEuyWZx/view

[21] IMAGINE-B5G, ‘D4.5 Intermediate Validation of B5G Core Technologies’, Dec. 2024. [Online]. Available: https://drive.google.com/file/d/13AJz1xw2J9GEKIGpsMiayhuncCeokpmf/view
[22] ORIGAMI, ‘ORIGAMI | Home’. Accessed: Apr. 19, 2024. [Online]. Available: https://sns-origami.eu/
[23] PRIVATEER, ‘PRIVATEER | Project | Privacy-first Security Enablers for 6G Networks | HORIZON’, CORDIS | European Commission. Accessed: Aug. 25, 2025. [Online]. Available: https://cordis.europa.eu/project/id/101096110
[24] Deterministic6G, ‘D1.1 Use Cases and Architecture Principles’, Jun. 2023.
[25] G. P. Sharma *et al.*, ‘Toward Deterministic Communications in 6G Networks: State of the Art, Open Challenges and the Way Forward’, *IEEE Access*, vol. 11, pp. 106898–106923, 2023, doi: 10.1109/ACCESS.2023.3316605.
[26] SAFE-6G, ‘D2.1 Definition of Technical Requirements for User-centric 6G Trustworthiness’, Aug. 2024. Accessed: Aug. 25, 2025. [Online]. Available: https://safe-6g.eu/wp-content/uploads/2024/10/D2.1_SAFE-6G_v1.0.pdf
[27] D. Tsolkas *et al.*, ‘Network & Service Management Advancements - Key frameworks and Interfaces towards open, Intelligent and reliable 6G networks’, Mar. 2025, doi: 10.5281/ZENODO.14234897.
[28] G. Callebaut *et al.*, ‘An Open Dataset Storage Standard for 6G Testbeds’, 2023, doi: 10.48550/ARXIV.2311.02662.
[29] P. Agddam, ‘6GTandem KPI/KVI’, presented at the Stream B/D Joint Workshop on KPIs and KVIs, May 16, 2023. [Online]. Available: https://smart-networks.europa.eu/wp-content/uploads/2024/05/6gtandem-stream-b-kpi-kvi-presentation.pdf
[30] 6G-TAMDEM, ‘6gTANDEM’. Accessed: Aug. 25, 2025. [Online]. Available: https://horizon-6gtandem.eu/
[31] W. Chen *et al.*, ‘5G-Advanced Toward 6G: Past, Present, and Future’, *IEEE Journal on Selected Areas in Communications*, vol. 41, no. 6, pp. 1592–1619, Jun. 2023, doi: 10.1109/JSAC.2023.3274037.
[32] PREDICT-6G, ‘D1.1 Analysis of use cases and system requirements’, Jul. 2023, Accessed: Aug. 25, 2025. [Online]. Available: https://zenodo.org/records/12167871
[33] 6G-LEADER, ‘D7.1 Prototype Plan and Evaluation Methodology.’, Aug. 2025.
[34] UN, ‘THE 17 GOALS | Sustainable Development’. Accessed: Aug. 25, 2025. [Online]. Available: https://sdgs.un.org/goals
[35] A. Sahin and R. Yang, ‘A Survey on Over-the-Air Computation’, 2022, *arXiv*. doi: 10.48550/ARXIV.2210.11350.
[36] C. Zhang, Y. Xie, H. Bai, B. Yu, W. Li, and Y. Gao, ‘A survey on federated learning’, *Knowledge-Based Systems*, vol. 216, p. 106775, Mar. 2021, doi: 10.1016/j.knosys.2021.106775.
[37] K. Yang, T. Jiang, Y. Shi, and Z. Ding, ‘Federated Learning via Over-the-Air Computation’, 2018, *arXiv*. doi: 10.48550/ARXIV.1812.11750.
[38] Z. Yang, M. Chen, K.-K. Wong, H. V. Poor, and S. Cui, ‘Federated Learning for 6G: Applications, Challenges, and Opportunities’, 2021, *arXiv*. doi: 10.48550/ARXIV.2101.01338.
[39] G. Zhu, J. Xu, K. Huang, and S. Cui, ‘Over-the-Air Computing for Wireless Data Aggregation in Massive IoT’, *IEEE Wireless Commun.*, vol. 28, no. 4, pp. 57–65, Aug. 2021, doi: 10.1109/MWC.011.2000467.
[40] H. Xing, O. Simeone, and S. Bi, ‘Federated Learning over Wireless Device-to-Device Networks: Algorithms and Convergence Analysis’, 2021, doi: 10.48550/ARXIV.2101.12704.
[41] M. Badi, C. B. Issaid, A. Elgabli, and M. Bennis, ‘Balancing Energy Efficiency and Distributional Robustness in Over-the-Air Federated Learning’, in *2024 IEEE International*

Conference on Machine Learning for Communication and Networking (ICMLCN), Stockholm, Sweden: IEEE, May 2024, pp. 195–200. doi: 10.1109/ICMLCN59089.2024.10624815.

[42] Z. Wang, Y. Zhao, Y. Zhou, Y. Shi, C. Jiang, and K. B. Letaief, 'Over-the-Air Computation for 6G: Foundations, Technologies, and Applications', *IEEE Internet Things J.*, vol. 11, no. 14, pp. 24634–24658, Jul. 2024, doi: 10.1109/JIOT.2024.3405448.

[43] K. Dong, S. A. Vorobyov, H. Yu, and T. Taleb, 'Beamforming Design for Integrated Sensing, Over-the-Air Computation, and Communication in Internet of Robotic Things', *IEEE Internet Things J.*, vol. 11, no. 20, pp. 32478–32489, Oct. 2024, doi: 10.1109/JIOT.2024.3433390.

[44] S. Ramnath, A. Javali, B. Narang, P. Mishra, and S. K. Routray, 'IoT based localization and tracking', in *2017 International Conference on IoT and Application (ICIOT)*, Nagapattinam, India: IEEE, May 2017, pp. 1–4. doi: 10.1109/ICIOTA.2017.8073629.

[45] S. S. Muhtasimul Hoque, M. H. Adeli, and A. Sahin, 'Chirp-Based Over-the-Air Computation for Long-Range Federated Edge Learning', in *2022 IEEE 33rd Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, Kyoto, Japan: IEEE, Sep. 2022, pp. 720–726. doi: 10.1109/PIMRC54779.2022.9978134.

[46] B. Doorgakant, T. P. Fowdur, and M. O. Akinsolu, 'End-to-End Power Models for 5G Radio Access Network Architectures with a Perspective on 6G', *Mathematics*, vol. 13, no. 3, p. 466, Jan. 2025, doi: 10.3390/math13030466.

[47] VIAVI Solutions, 'What is 5G Energy Consumption?' Accessed: Aug. 25, 2025. [Online]. Available: https://www.viavisolutions.com/en-uk/resources/learning-center/what-5g-energy-consumption

[48] S. Wesemann, J. Du, and H. Viswanathan, 'Energy Efficient Extreme MIMO: Design Goals and Directions', 2023, doi: 10.48550/ARXIV.2301.01119.

[49] O-RAN, 'WG1 TS "Use Cases Detailed Specification" R004 v17.00', Alfter, Germany, O-RAN.WG1.TS.Use-Cases-Detailed-Specification-R004-v16.00, Jun. 2025. [Online]. Available: https://specifications.o-ran.org/download?id=865

[50] O-RAN, 'WG1 TR "Use Cases Analysis Report" R004 v16.00', Alfter, Germany, O-RAN.WG1.TR.Use-Cases-Analysis-Report-R004-v16.00, Feb. 2025.

[51] V. Jamali, A. M. Tulino, G. Fischer, R. R. Müller, and R. Schober, 'Intelligent Surface-Aided Transmitter Architectures for Millimeter-Wave Ultra Massive MIMO Systems', *IEEE Open Journal of the Communications Society*, vol. 2, pp. 144–167, 2021, doi: 10.1109/OJCOMS.2020.3048063.

[52] 3GPP, 'Directory Listing /ftp/workshop/2025-03-10_3GPP_6G_WS/Docs'. Accessed: Aug. 25, 2025. [Online]. Available: https://www.3gpp.org/ftp/workshop/2025-03-10_3GPP_6G_WS/Docs

[53] '3GPP workshop on 6G'. Accessed: Sep. 02, 2025. [Online]. Available: https://www.3gpp.org/news-events/3gpp-news/6gworkshop-2025

[54] A. M. Elbir, K. V. Mishra, M. R. B. Shankar, and B. Ottersten, 'A Family of Deep Learning Architectures for Channel Estimation and Hybrid Beamforming in Multi-Carrier mm-Wave Massive MIMO', *IEEE Transactions on Cognitive Communications and Networking*, vol. 8, no. 2, pp. 642–656, Jun. 2022, doi: 10.1109/TCCN.2021.3132609.

[55] K. Ma, D. He, H. Sun, Z. Wang, and S. Chen, 'Deep Learning Assisted Calibrated Beam Training for Millimeter-Wave Communication Systems', *IEEE Transactions on Communications*, vol. 69, no. 10, pp. 6706–6721, Oct. 2021, doi: 10.1109/TCOMM.2021.3098683.

[56] F. Sohrabi, Z. Chen, and W. Yu, 'Deep Active Learning Approach to Adaptive Beamforming for mmWave Initial Alignment', *IEEE Journal on Selected Areas in Communications*, vol. 39, no. 8, pp. 2347–2360, Aug. 2021, doi: 10.1109/JSAC.2021.3087234.

[57] H. He, C.-K. Wen, S. Jin, and G. Y. Li, 'Deep Learning-Based Channel Estimation for Beamspace mmWave Massive MIMO Systems', *IEEE Wireless Communications Letters*, vol. 7, no. 5, pp. 852–855, Oct. 2018, doi: 10.1109/LWC.2018.2832128.
[58] J. Gao, C. Zhong, G. Y. Li, J. B. Soriaga, and A. Behboodi, 'Deep Learning-Based Channel Estimation for Wideband Hybrid MmWave Massive MIMO', *IEEE Transactions on Communications*, vol. 71, no. 6, pp. 3679–3693, Jun. 2023, doi: 10.1109/TCOMM.2023.3258484.
[59] X. Ma, Z. Gao, F. Gao, and M. Di Renzo, 'Model-Driven Deep Learning Based Channel Estimation and Feedback for Millimeter-Wave Massive Hybrid MIMO Systems', *IEEE Journal on Selected Areas in Communications*, vol. 39, no. 8, pp. 2388–2406, Aug. 2021, doi: 10.1109/JSAC.2021.3087269.
[60] W. Jiang and H. D. Schotten, 'Deep Learning for Fading Channel Prediction', *IEEE Open Journal of the Communications Society*, vol. 1, pp. 320–332, 2020, doi: 10.1109/OJCOMS.2020.2982513.
[61] H. Jiang, M. Cui, D. W. K. Ng, and L. Dai, 'Accurate Channel Prediction Based on Transformer: Making Mobility Negligible', *IEEE Journal on Selected Areas in Communications*, vol. 40, no. 9, pp. 2717–2732, Sep. 2022, doi: 10.1109/JSAC.2022.3191334.
[62] J.-H. Lee, J. Lee, and A. F. Molisch, 'Generative vs. Predictive Models in Massive MIMO Channel Prediction', in *2024 58th Asilomar Conference on Signals, Systems, and Computers*, Oct. 2024, pp. 1642–1646. doi: 10.1109/IEEECONF60004.2024.10943090.
[63] Y. Wu, Y. Li, R. Liang, G. Zheng, and X. Wang, 'CGAN-Based Data Augmentation for Enhanced Channel Prediction in Massive MIMO Under Subway Tunnels', *IEEE Communications Letters*, vol. 29, no. 6, pp. 1255–1259, Jun. 2025, doi: 10.1109/LCOMM.2025.3558561.
[64] Z. Lu, C. Zhong, and M. C. Gursoy, 'Dynamic Channel Access and Power Control in Wireless Interference Networks via Multi-Agent Deep Reinforcement Learning', *IEEE Transactions on Vehicular Technology*, vol. 71, no. 2, pp. 1588–1601, Feb. 2022, doi: 10.1109/TVT.2021.3131534.
[65] F. Meng, P. Chen, L. Wu, and J. Cheng, 'Power Allocation in Multi-User Cellular Networks: Deep Reinforcement Learning Approaches', *IEEE Transactions on Wireless Communications*, vol. 19, no. 10, pp. 6255–6267, Oct. 2020, doi: 10.1109/TWC.2020.3001736.
[66] H. Zhang, H. Zhang, K. Long, and G. K. Karagiannidis, 'Deep Learning Based Radio Resource Management in NOMA Networks: User Association, Subchannel and Power Allocation', *IEEE Transactions on Network Science and Engineering*, vol. 7, no. 4, pp. 2406–2415, Oct. 2020, doi: 10.1109/TNSE.2020.3004333.
[67] W. Lee and R. Schober, 'Deep Learning-Based Resource Allocation for Device-to-Device Communication', *IEEE Transactions on Wireless Communications*, vol. 21, no. 7, pp. 5235–5250, Jul. 2022, doi: 10.1109/TWC.2021.3138733.
[68] S. L. Shah, N. H. Mahmood, and M. Latva-aho, 'Interference Prediction Using Gaussian Process Regression and Management Framework for Critical Services in Local 6G Networks', in *2025 IEEE Wireless Communications and Networking Conference (WCNC)*, Mar. 2025, pp. 1–6. doi: 10.1109/WCNC61545.2025.10978635.
[69] S. Huang, Y. Ye, and M. Xiao, 'Learning-Based Hybrid Beamforming Design for Full-Duplex Millimeter Wave Systems', *IEEE Transactions on Cognitive Communications and Networking*, vol. 7, no. 1, pp. 120–132, Mar. 2021, doi: 10.1109/TCCN.2020.3019604.
[70] J. Chen *et al.*, 'Hybrid Beamforming/Combining for Millimeter Wave MIMO: A Machine Learning Approach', *IEEE Transactions on Vehicular Technology*, vol. 69, no. 10, pp. 11353–11368, Oct. 2020, doi: 10.1109/TVT.2020.3009746.

[71] W. Wang and W. Zhang, 'Intelligent Reflecting Surface Configurations for Smart Radio Using Deep Reinforcement Learning', *IEEE Journal on Selected Areas in Communications*, vol. 40, no. 8, pp. 2335–2346, Aug. 2022, doi: 10.1109/JSAC.2022.3180787.

[72] S. Gong, J. Lin, B. Ding, D. Niyato, D. I. Kim, and M. Guizani, 'When Optimization Meets Machine Learning: The Case of IRS-Assisted Wireless Networks', *IEEE Network*, vol. 36, no. 2, pp. 190–198, Mar. 2022, doi: 10.1109/MNET.211.2100386.

[73] A. M. Girgis, J. Park, M. Bennis, and M. Debbah, 'Predictive Control and Communication Co-Design via Two-Way Gaussian Process Regression and AoI-Aware Scheduling', *IEEE Transactions on Communications*, vol. 69, no. 10, pp. 7077–7093, Oct. 2021, doi: 10.1109/TCOMM.2021.3099156.

[74] X. Cao, G. Zhu, J. Xu, and S. Cui, 'Transmission Power Control for Over-the-Air Federated Averaging at Network Edge', 2021, *arXiv*. doi: 10.48550/ARXIV.2111.05719.

[75] W. Fang, Y. Jiang, Y. Shi, Y. Zhou, W. Chen, and K. B. Letaief, 'Over-the-Air Computation via Reconfigurable Intelligent Surface', *IEEE Trans. Commun.*, vol. 69, no. 12, pp. 8612–8626, Dec. 2021, doi: 10.1109/TCOMM.2021.3114791.

[76] W. Zhang, J. Xu, W. Xu, X. You, and W. Fu, 'Worst-case Design for RIS-aided Over-the-air Computation with Imperfect CSI', 2022, *arXiv*. doi: 10.48550/ARXIV.2206.06936.

[77] W. Guo, R. Li, C. Huang, X. Qin, K. Shen, and W. Zhang, 'Joint Device Selection and Power Control for Wireless Federated Learning', *IEEE J. Select. Areas Commun.*, vol. 40, no. 8, pp. 2395–2410, Aug. 2022, doi: 10.1109/JSAC.2022.3180807.

[78] C. Xu, S. Liu, Z. Yang, Y. Huang, and K.-K. Wong, 'Learning Rate Optimization for Federated Learning Exploiting Over-the-air Computation', 2021, doi: 10.48550/ARXIV.2102.02946.

[79] H. Hellström, V. Fodor, and C. Fischione, 'Federated Learning Over-the-Air by Retransmissions', *IEEE Trans. Wireless Commun.*, vol. 22, no. 12, pp. 9143–9156, Dec. 2023, doi: 10.1109/TWC.2023.3268742.

[80] Z. Wang, Y. Zhou, Y. Shi, and W. Zhuang, 'Interference Management for Over-the-Air Federated Learning in Multi-Cell Wireless Networks', 2022, doi: 10.48550/ARXIV.2206.02398.

[81] L. Dai, B. Wang, Y. Yuan, S. Han, I. Chih-lin, and Z. Wang, 'Non-orthogonal multiple access for 5G: solutions, challenges, opportunities, and future research trends', *IEEE Commun. Mag.*, vol. 53, no. 9, pp. 74–81, Sep. 2015, doi: 10.1109/MCOM.2015.7263349.

[82] M. S. Ali, H. Tabassum, and E. Hossain, 'Dynamic User Clustering and Power Allocation for Uplink and Downlink Non-Orthogonal Multiple Access (NOMA) Systems', *IEEE Access*, pp. 1–1, 2016, doi: 10.1109/ACCESS.2016.2604821.

[83] J. Zhang, H. Zheng, Z. Chen, X. Chen, and G. Min, 'Device Access, Subchannel Division, and Transmission Power Allocation for NOMA-Enabled IoT Systems', *IEEE Internet Things J.*, vol. 10, no. 19, pp. 17047–17057, Oct. 2023, doi: 10.1109/JIOT.2023.3273306.

[84] H. Joudeh and B. Clerckx, 'Robust Transmission in Downlink Multiuser MISO Systems: A Rate-Splitting Approach', *IEEE Trans. Signal Process.*, vol. 64, no. 23, pp. 6227–6242, Dec. 2016, doi: 10.1109/TSP.2016.2591501.

[85] Z. Yang, M. Chen, W. Saad, W. Xu, and M. Shikh-Bahaei, 'Sum-Rate Maximization of Uplink Rate Splitting Multiple Access (RSMA) Communication', in *2019 IEEE Global Communications Conference (GLOBECOM)*, Waikoloa, HI, USA: IEEE, Dec. 2019, pp. 1–6. doi: 10.1109/GLOBECOM38437.2019.9013344.

[86] M. B. Shahab, R. Abbas, M. Shirvanimoghaddam, and S. J. Johnson, 'Grant-Free Non-Orthogonal Multiple Access for IoT: A Survey', *IEEE Commun. Surv. Tutorials*, vol. 22, no. 3, pp. 1805–1838, 2020, doi: 10.1109/COMST.2020.2996032.

[87] M. Fayaz, W. Yi, Y. Liu, S. Thayaparan, and A. Nallanathan, 'Toward Autonomous Power Control in Semi-Grant-Free NOMA Systems: A Power Pool-Based Approach', *IEEE Trans. Commun.*, vol. 72, no. 6, pp. 3273–3289, Jun. 2024, doi: 10.1109/TCOMM.2024.3361535.

[88] N. Díaz-Rodríguez, J. Del Ser, M. Coeckelbergh, M. López De Prado, E. Herrera-Viedma, and F. Herrera, ‘Connecting the dots in trustworthy Artificial Intelligence: From AI principles, ethics, and key requirements to responsible AI systems and regulation’, *Information Fusion*, vol. 99, p. 101896, Nov. 2023, doi: 10.1016/j.inffus.2023.101896.
[89] M. Huh, B. Cheung, T. Wang, and P. Isola, ‘Position: The Platonic Representation Hypothesis’, in *Proceedings of the 41st International Conference on Machine Learning*, PMLR, Jul. 2024, pp. 20617–20642. Accessed: Aug. 25, 2025. [Online]. Available: https://proceedings.mlr.press/v235/huh24a.html
[90] T. Kaufmann, P. Weng, V. Bengs, and E. Hüllermeier, ‘A Survey of Reinforcement Learning from Human Feedback’, Apr. 30, 2024, *arXiv*: arXiv:2312.14925. doi: 10.48550/arXiv.2312.14925.
[91] M. Kountouris and N. Pappas, ‘Semantics-Empowered Communication for Networked Intelligent Systems’, *IEEE Commun. Mag.*, vol. 59, no. 6, pp. 96–102, Jun. 2021, doi: 10.1109/MCOM.001.2000604.
[92] P. Popovski *et al.*, ‘A Perspective on Time Toward Wireless 6G’, *Proc. IEEE*, vol. 110, no. 8, pp. 1116–1146, Aug. 2022, doi: 10.1109/JPROC.2022.3190205.
[93] D. Gunduz *et al.*, ‘Beyond Transmitting Bits: Context, Semantics, and Task-Oriented Communications’, *IEEE J. Select. Areas Commun.*, vol. 41, no. 1, pp. 5–41, Jan. 2023, doi: 10.1109/JSAC.2022.3223408.
[94] S. Kaul, R. Yates, and M. Gruteser, ‘Real-time status: How often should one update?’, in *2012 Proceedings IEEE INFOCOM*, Mar. 2012, pp. 2731–2735. doi: 10.1109/INFCOM.2012.6195689.
[95] R. D. Yates, Y. Sun, D. R. Brown, S. K. Kaul, E. Modiano, and S. Ulukus, ‘Age of Information: An Introduction and Survey’, *IEEE Journal on Selected Areas in Communications*, vol. 39, no. 5, pp. 1183–1210, May 2021, doi: 10.1109/JSAC.2021.3065072.
[96] N. Pappas, M. A. Abd-Elmagid, B. Zhou, W. Saad, and H. S. Dhillon, *Age of Information: Foundations and Applications*. Cambridge University Press, 2023.
[97] A. Kosta, N. Pappas, A. Ephremides, and V. Angelakis, ‘The Cost of Delay in Status Updates and Their Value: Non-Linear Ageing’, *IEEE Transactions on Communications*, vol. 68, no. 8, pp. 4905–4918, Aug. 2020, doi: 10.1109/TCOMM.2020.2988013.
[98] M. E. Ildiz, O. T. Yavascan, E. Uysal, and O. T. Kartal, ‘Pull or Wait: How to Optimize Query Age of Information’, *IEEE Journal on Selected Areas in Information Theory*, vol. 4, pp. 794–807, 2023, doi: 10.1109/JSAIT.2023.3346308.
[99] A. Maatouk, M. Assaad, and A. Ephremides, ‘The Age of Incorrect Information: An Enabler of Semantics-Empowered Communication’, *IEEE Transactions on Wireless Communications*, vol. 22, no. 4, pp. 2621–2635, Apr. 2023, doi: 10.1109/TWC.2022.3213227.
[100] Y. Sun, Y. Polyanskiy, and E. Uysal, ‘Sampling of the Wiener Process for Remote Estimation Over a Channel With Random Delay’, *IEEE Transactions on Information Theory*, vol. 66, no. 2, pp. 1118–1135, Feb. 2020, doi: 10.1109/TIT.2019.2937336.
[101] O. Ayan, M. Vilgelm, M. Klügel, S. Hirche, and W. Kellerer, ‘Age-of-information vs. value-of-information scheduling for cellular networked control systems’, in *Proceedings of the 10th ACM/IEEE International Conference on Cyber-Physical Systems*, in ICCPS ’19. New York, NY, USA: Association for Computing Machinery, Apr. 2019, pp. 109–117. doi: 10.1145/3302509.3311050.
[102] V. N. Swamy, P. Rigge, G. Ranade, B. Nikolić, and A. Sahai, ‘Wireless Channel Dynamics and Robustness for Ultra-Reliable Low-Latency Communications’, *IEEE Journal on Selected Areas in Communications*, vol. 37, no. 4, pp. 705–720, Apr. 2019, doi: 10.1109/JSAC.2019.2900784.

[103] G. Ranade and A. Sahai, ‘Control Capacity’, *IEEE Transactions on Information Theory*, vol. 65, no. 1, pp. 235–254, Jan. 2019, doi: 10.1109/TIT.2018.2868929.
[104] FAIRsharing Team, ‘FAIRsharing record for: Quantities, Units, Dimensions and Types’. FAIRsharing, 2015. doi: 10.25504/FAIRSHARING.D3PQW7.
[105] E. Hechler, M. Weihrauch, and Y. Wu, ‘Data Fabric Architecture Patterns’, in *Data Fabric and Data Mesh Approaches with AI*, Berkeley, CA: Apress, 2023, pp. 231–255. doi: 10.1007/978-1-4842-9253-2_10.
[106] J. Barrasa and J. Webber, *Building Knowledge Graphs. A Practitioner’s Guide*, First. 1005 Gravenstein Highway North, Sebastopol, CA 95472.: O’Reilly Media, Inc, 2023. [Online]. Available: https://go.neo4j.com/rs/710-RRC-335/images/Building-Knowledge-Graphs-Practitioner's-Guide-OReilly-book.pdf
[107] S. Martin, B. Szekely, and D. Allemang, The rise of the knowledge graph: Toward modern data integration and the data fabric architecture. O’Reilly Media, Incorporated, 2021., *The Rise of The Knowledge Graph*, First. O’Reilly Media, Inc., 2021. Accessed: Aug. 26, 2025. [Online]. Available: https://www.scribd.com/document/750144791/The-Rise-of-the-Knowledge-Graph
[108] Gergana Petkova, ‘What Is Data Fabric?’, Ontotext. Accessed: Aug. 26, 2025. [Online]. Available: https://www.ontotext.com/knowledgehub/fundamentals/what-is-data-fabric/
[109] A. Hogan *et al.*, ‘Knowledge Graphs’, *ACM Comput. Surv.*, vol. 54, no. 4, pp. 1–37, May 2022, doi: 10.1145/3447772.
[110] T. Berners-Lee, J. Hendler, and O. Lassila, ‘The Semantic Web’, *Sci Am*, vol. 284, no. 5, pp. 34–43, May 2001, doi: 10.1038/scientificamerican0501-34.
[111] Marco Grassi, Mario Scrocca, Alessio Carenini, Marco Comerio, and Irene Celino, ‘Composable Semantic Data Transformation Pipelines with Chimera’, Jun. 2023, doi: 10.5281/ZENODO.8020088.
[112] P. Li and A. Aijaz, ‘Open RAN meets Semantic Communications: A Synergistic Match for Open, Intelligent, and Knowledge-Driven 6G’, Oct. 15, 2023, *arXiv*: arXiv:2310.09951. doi: 10.48550/arXiv.2310.09951.
[113] E. C. Strinati *et al.*, ‘Goal-Oriented and Semantic Communication in 6G AI-Native Networks: The 6G-GOALS Approach’, in *2024 Joint European Conference on Networks and Communications & 6G Summit (EuCNC/6G Summit)*, Antwerp, Belgium: IEEE, Jun. 2024, pp. 1–6. doi: 10.1109/EuCNC/6GSummit60053.2024.10597087.
[114] H. Zhou, Y. Deng, X. Liu, N. Pappas, and A. Nallanathan, ‘Goal-Oriented Semantic Communications for 6G Networks’, Apr. 06, 2024, *arXiv*: arXiv:2210.09372. doi: 10.48550/arXiv.2210.09372.
[115] M. Shokrnezhad, H. Mazandarani, T. Taleb, J. Song, and R. Li, ‘Semantic Revolution From Communications to Orchestration for 6G: Challenges, Enablers, and Research Directions’, *IEEE Network*, vol. 38, no. 6, pp. 63–71, Nov. 2024, doi: 10.1109/MNET.2024.3422348.
[116] G. Kakkavas, D. Gkatzioura, V. Karyotis, and S. Papavassiliou, ‘A Review of Advanced Algebraic Approaches Enabling Network Tomography for Future Network Infrastructures’, *Future Internet*, vol. 12, no. 2, p. 20, Jan. 2020, doi: 10.3390/fi12020020.
[117] G. Kakkavas, A. Stamou, V. Karyotis, and S. Papavassiliou, ‘Network Tomography for Efficient Monitoring in SDN-Enabled 5G Networks and Beyond: Challenges and Opportunities’, *IEEE Commun. Mag.*, vol. 59, no. 3, pp. 70–76, Mar. 2021, doi: 10.1109/MCOM.001.2000458.
[118] A. Madnaik, N. C. Matson, and K. Sundaresan, ‘Scalable Network Tomography for Dynamic Spectrum Access’, in *IEEE INFOCOM 2024 - IEEE Conference on Computer Communications*, Vancouver, BC, Canada: IEEE, May 2024, pp. 2209–2218. doi: 10.1109/INFOCOM52122.2024.10621172.

[119] W. Kim, Y. Ahn, J. Kim, and B. Shim, 'Towards deep learning-aided wireless channel estimation and channel state information feedback for 6G', *J. Commun. Netw.*, vol. 25, no. 1, pp. 61–75, Feb. 2023, doi: 10.23919/JCN.2022.000037.

[120] G. Kakkavas, N. Fryganiotis, V. Karyotis, and S. Papavassiliou, 'Generative Deep Learning Techniques for Traffic Matrix Estimation From Link Load Measurements', *IEEE Open J. Commun. Soc.*, vol. 5, pp. 1029–1046, 2024, doi: 10.1109/OJCOMS.2024.3358740.

[121] G. Kakkavas, P. Maratos, V. Karyotis, and S. Papavassiliou, 'Traffic Matrix Estimation Using Invertible Neural Networks', in *2024 International Conference on Software, Telecommunications and Computer Networks (SoftCOM)*, Split, Croatia: IEEE, Sep. 2024, pp. 1–7. doi: 10.23919/SoftCOM62040.2024.10721829.

[122] Y. Wang, C. Yang, T. Li, Y. Ouyang, X. Mi, and Y. Song, 'A Survey on Intent-Driven End-to-End 6G Mobile Communication System', *IEEE Commun. Surv. Tutorials*, pp. 1–1, 2025, doi: 10.1109/COMST.2025.3575041.

[123] Y. Ouyang, C. Li, J. Zhang, X. Zhao, and C. Yang, 'Intent-Driven 6G End-to-End Network Orchestration', in *IEEE INFOCOM 2024 - IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*, Vancouver, BC, Canada: IEEE, May 2024, pp. 1–2. doi: 10.1109/INFOCOMWKSHPS61880.2024.10620891.

[124] B. Ojaghi, R. Vilalta, and R. Muñoz, 'Intent-Based Network Resource Slicing in 6G', in *2024 15th International Conference on Network of the Future (NoF)*, Castelldefels, Spain: IEEE, Oct. 2024, pp. 31–37. doi: 10.1109/NoF62948.2024.10741513.

[125] M. A. Habib *et al.*, 'Harnessing the Power of LLMs, Informers and Decision Transformers for Intent-driven RAN Management in 6G', 2025, *arXiv*. doi: 10.48550/ARXIV.2505.01841.

[126] D. Brodimas *et al.*, 'Towards Intent-based Network Management for the 6G System adopting Multimodal Generative AI', in *2024 Joint European Conference on Networks and Communications & 6G Summit (EuCNC/6G Summit)*, Antwerp, Belgium: IEEE, Jun. 2024, pp. 848–853. doi: 10.1109/EuCNC/6GSummit60053.2024.10597022.

[127] J. Zhang *et al.*, 'Intent-Driven Closed-Loop Control and Management Framework for 6G Open RAN', *IEEE Internet Things J.*, vol. 11, no. 4, pp. 6314–6327, Feb. 2024, doi: 10.1109/JIOT.2023.3312795.

[128] P. Alemany *et al.*, 'Defining Intent-Based Service Management Automation for 6G Multi-Stakeholders Scenarios', *IEEE Open J. Commun. Soc.*, vol. 6, pp. 2373–2396, 2025, doi: 10.1109/OJCOMS.2025.3554250.

[129] C. Vassilakis, N. Fryganiotis, P. Maratos, A. Zafeiropoulos, E. Stai, and S. Papavassiliou, 'Intent Lifecycle Management over Programmable Infrastructure in the Computing Continuum', in *Proceedings of the 2nd International Workshop on MetaOS for the Cloud-Edge-IoT Continuum*, Rotterdam Netherlands: ACM, Mar. 2025, pp. 46–52. doi: 10.1145/3721889.3721928.

[130] Matt Rutkowski, Chris Lauwers, Claude Noshpitz, and Calin Curescu, *TOSCA Simple Profile in YAML Version 1.3*, Feb. 26, 2020. Accessed: Aug. 26, 2025. [Online]. Available: https://docs.oasis-open.org/tosca/TOSCA-Simple-Profile-YAML/v1.3/TOSCA-Simple-Profile-YAML-v1.3.html

[131] K.-K. Wong, A. Shojaeifard, K.-F. Tong, and Y. Zhang, 'Fluid Antenna Systems', *IEEE Trans. Wireless Commun.*, vol. 20, no. 3, pp. 1950–1962, Mar. 2021, doi: 10.1109/TWC.2020.3037595.

[132] K.-K. Wong and K.-F. Tong, 'Fluid Antenna Multiple Access', *IEEE Trans. Wireless Commun.*, vol. 21, no. 7, pp. 4801–4815, Jul. 2022, doi: 10.1109/TWC.2021.3133410.

[133] Y. Shen, K.-F. Tong, and K.-K. Wong, 'Radiation Pattern Diversified Double-Fluid-Channel Surface-Wave Antenna for Mobile Communications', in *2022 IEEE-APS Topical Conference on Antennas and Propagation in Wireless Communications (APWC)*, Cape Town, South Africa: IEEE, Sep. 2022, pp. 085–088. doi: 10.1109/APWC49427.2022.9899924.

[134] Y. Liu *et al.*, 'Reconfigurable Intelligent Surfaces: Principles and Opportunities', *IEEE Communications Surveys & Tutorials*, vol. 23, no. 3, pp. 1546–1577, 2021, doi: 10.1109/COMST.2021.3077737.

[135] J. Kim and M. Andrews, 'Learning-Based Adaptive User Selection in Millimeter Wave Hybrid Beamforming Systems', in *ICC 2023 - IEEE International Conference on Communications*, Rome, Italy: IEEE, May 2023, pp. 5645–5650. doi: 10.1109/ICC45041.2023.10279582.

[136] T. Kobal, F. Durand, R. Koblitz, and M. Andrews, 'Post-Computing Analog Beams after User Selection in a Hybrid Beamforming System', in *2024 Joint European Conference on Networks and Communications & 6G Summit (EuCNC/6G Summit)*, Antwerp, Belgium: IEEE, Jun. 2024, pp. 457–462. doi: 10.1109/EuCNC/6GSummit60053.2024.10597109.

[137] P. Agheli, T. Kobal, F. Durand, and M. Andrews, 'Learning-Based Multiuser Scheduling in Mimo-Ofdm Systems With Hybrid Beamforming', in *2025 Joint European Conference on Networks and Communications & 6G Summit (EuCNC/6G Summit)*, Poznan, Poland: IEEE, Jun. 2025, pp. 1–6. doi: 10.1109/EuCNC/6GSummit63408.2025.11037174.

[138] Balaji Raghothaman *et al.*, 'Research Report on Native and Cross-domain AI: State of the art and future outlook', O-RAN next Generation Research Group (nGRG), RR-2023-03, Sep. 2023. Accessed: Aug. 26, 2025. [Online]. Available: https://mediastorage.o-ran.org/ngrg-rr/nGRG-RR-2023-03-Research-Report-on-Native-and-Cross-domain-AI-v1_1.pdf

[139] Ziqi Chen *et al.*, 'Research Report on Generative AI Use Cases and Requirements on 6G Network', O-RAN next Generation Research Group (nGRG), RS-2025-02, Jun. 2025. Accessed: Aug. 26, 2025. [Online]. Available: https://mediastorage.o-ran.org/ngrg-rr/nGRG-RR-2025-02-GenAI%20in%20Network.pdf

[140] P. Gajjar and V. K. Shah, 'ORANSight-2.0: Foundational LLMs for O-RAN', Jul. 22, 2025, *arXiv*: arXiv:2503.05200. doi: 10.48550/arXiv.2503.05200.

[141] S. Roy, H. Chergui, A. Ksentini, and C. Verikoukis, 'Federated Machine Reasoning for Resource Provisioning in 6G O-RAN', Jun. 10, 2024, *arXiv*: arXiv:2406.06128. doi: 10.48550/arXiv.2406.06128.

[142] B. Brik *et al.*, 'Explainable AI in 6G O-RAN: A Tutorial and Survey on Architecture, Use Cases, Challenges, and Future Research', *IEEE Commun. Surv. Tutorials*, pp. 1–1, 2024, doi: 10.1109/COMST.2024.3510543.

[143] C. Fiandrino, L. Bonati, S. D'Oro, M. Polese, T. Melodia, and J. Widmer, 'EXPLORA: AI/ML EXPLainability for the Open RAN', *Proc. ACM Netw.*, vol. 1, no. CoNEXT3, pp. 1–26, Nov. 2023, doi: 10.1145/3629141.

[144] M. A. Habibi, B. Han, M. Saimler, I. L. Pavon, and H. D. Schotten, 'Towards an AI/ML-driven SMO Framework in O-RAN: Scenarios, Solutions, and Challenges', Sep. 08, 2024, *arXiv*: arXiv:2409.05092. doi: 10.48550/arXiv.2409.05092.

[145] M. A. Habibi *et al.*, 'Toward an Open, Intelligent, and End-to-End Architectural Framework for Network Slicing in 6G Communication Systems', *IEEE Open J. Commun. Soc.*, vol. 4, pp. 1615–1658, 2023, doi: 10.1109/OJCOMS.2023.3294445.

[146] M. M. H. Qazzaz, Ł. Kułacz, A. Kliks, S. A. Zaidi, M. Dryjanski, and D. McLernon, 'Machine Learning-based xApp for Dynamic Resource Allocation in O-RAN Networks', in *2024 IEEE International Conference on Machine Learning for Communication and Networking (ICMLCN)*, Stockholm, Sweden: IEEE, May 2024, pp. 492–497. doi: 10.1109/ICMLCN59089.2024.10625184.

[147] C. Puligheddu, J. Ashdown, C. F. Chiasserini, and F. Restuccia, 'SEM-O-RAN: Semantic O-RAN Slicing for Mobile Edge Offloading of Computer Vision Tasks', *IEEE Trans. on Mobile Comput.*, vol. 23, no. 7, pp. 7785–7800, Jul. 2024, doi: 10.1109/TMC.2023.3339056.

[148] Y. Sun, L. Zhang, L. Guo, J. Li, D. Niyato, and Y. Fang, 'S-RAN: Semantic-Aware Radio Access Networks', *IEEE Commun. Mag.*, vol. 63, no. 4, pp. 207–213, Apr. 2025, doi: 10.1109/MCOM.004.2400105.
[149] A. Masaracchia *et al.*, 'Toward 6G-Enabled URLLCs: Digital Twin, Open Ran, and Semantic Communications', *IEEE Comm. Stand. Mag.*, vol. 9, no. 1, pp. 13–20, Mar. 2025, doi: 10.1109/MCOMSTD.0001.2300054.
[150] M. Polese, L. Bonati, S. D'Oro, S. Basagni, and T. Melodia, 'Understanding O-RAN: Architecture, Interfaces, Algorithms, Security, and Research Challenges', *IEEE Commun. Surv. Tutorials*, vol. 25, no. 2, pp. 1376–1411, 2023, doi: 10.1109/COMST.2023.3239220.
[151] O-RAN, 'WG4 TS "O-RAN Fronthaul control, user and synchronization plane specification" R004 v07.00', Alfter, Germany, ORAN-WG4.CUS.0-v07.00, Jul. 2021.
[152] O-RAN, 'WG3 TS "O-RAN near-RT RAN intelligent controller near-RT RIC architecture" R004 v02.00', Alfter, Germany, O-RAN.WG3.RICARCH-v02.00, Mar. 2021.
[153] O-RAN, 'WG2 TS "O-RAN non-RT RIC architecture" R004 v01.00', Alfter, Germany, O-RAN.WG2.Non-RT-RIC-ARCH-TS-v01.00, Jul. 2021.
[154] O-RAN, 'WG1 TS "Use Cases and Overall Architecture: O-RAN Architecture Description" R004 v13.00', Alfter, Germany, O-RAN.WG1.TS.OAD-R004-v13.00, Feb. 2025.
[155] A. Bazzi, R. Bomfin, M. Mezzavilla, S. Rangan, T. Rappaport, and M. Chafii, 'Upper Mid-Band Spectrum for 6G: Vision, Opportunity and Challenges', 2025, *arXiv*. doi: 10.48550/ARXIV.2502.17914.
[156] M. Mohsin, J. M. Batalla, E. Pallis, G. Mastorakis, E. K. Markakis, and C. X. Mavromoustakis, 'On Analyzing Beamforming Implementation in O-RAN 5G', *Electronics*, vol. 10, no. 17, p. 2162, Sep. 2021, doi: 10.3390/electronics10172162.
[157] Salvatore D'Oro *et al.*, 'Research Report on dApps for Real-Time RAN Control: Use Cases and Requirements', O-RAN next Generation Research Group (nGRG), RR-2024-10, Oct. 2024. Accessed: Aug. 26, 2025. [Online]. Available: https://mediastorage.o-ran.org/ngrg-rr/nGRG-RR-2024-10-dApp%20use%20cases%20and%20requirements.pdf
[158] S. D'Oro, M. Polese, L. Bonati, H. Cheng, and T. Melodia, 'dApps: Distributed Applications for Real-Time Inference and Control in O-RAN', *IEEE Commun. Mag.*, vol. 60, no. 11, pp. 52–58, Nov. 2022, doi: 10.1109/MCOM.002.2200079.
[159] A. Lacava *et al.*, 'dApps: Enabling Real-Time AI-Based Open RAN Control', Jan. 27, 2025, *arXiv*: arXiv:2501.16502. doi: 10.48550/arXiv.2501.16502.
[160] TERRAMETA, 'D2.1 Requirements use cases and scenario specifications', Jun. 2023. Accessed: Aug. 26, 2025. [Online]. Available: https://terrameta-project.eu/wp-content/uploads/2023/09/TERRAMETA-Deliverable-D2.1-Requirements-use-cases-and-scenario-specifications.pdf
[161] TERRAMETA, 'D2.2 Network architecture and updated requirements and scenarios', Jul. 2024. Accessed: Aug. 26, 2025. [Online]. Available: https://terrameta-project.eu/wp-content/uploads/2024/09/TERRAMETA-Deliverable-D2.2-Network-architecture-and-updated-requirements-and-scenarios.pdf
[162] B. K. Jung, V. V. Elesina, S. Matos, R. D'Errico, and T. Kürner, 'Initial Assessment of THz Indoor Channel with Passive Reflective Intelligent Surfaces', in *2024 International Symposium on Antennas and Propagation (ISAP)*, Incheon, Korea, Republic of: IEEE, Nov. 2024, pp. 1–2. doi: 10.1109/ISAP62502.2024.10846784.
[163] 'BeGREEN', BeGREEN. Accessed: May 24, 2024. [Online]. Available: https://www.sns-begreen.com/
[164] M. Catalan-Cid, G. Castellanos, D. Reiss, and J. Armstrong, 'Demo: BeGREEN Intelligence Plane for AI-driven Energy Efficient O-RAN Management', presented at the IEEE INFOCOM 2025 Demo, 2025.,
[165] BeGREEN, 'D4.1 State-of-the-Art Review and Initial Definition of BeGREEN O-RAN Intelligent Plane, and AI/ML Algorithms for NFV User-Plane and Edge Service Control

Energy Efficiency Optimization'. Accessed: Aug. 26, 2025. [Online]. Available: https://s3.eu-west-1.amazonaws.com/cdn.webfactore.co.uk/sr_1852113.pdf?t=1710000706
[166] BeGREEN, 'D2.2 Evolved Architecture and Power Enhancement Mechanisms', Jul. 2024. Accessed: Aug. 26, 2025. [Online]. Available: https://s3.eu-west-1.amazonaws.com/cdn.webfactore.co.uk/4ce86c6d-91b2-4ce5-be06-6c4490c20185.pdf?t=1734001435
[167] '74 New or Updated O-RAN Technical Documents Released since July 2024'. Accessed: Sep. 22, 2025. [Online]. Available: https://www.o-ran.org/blog/74-new-or-updated-o-ran-technical-documents-released-since-july-2024
[168] C. Adamczyk, 'Challenges for Conflict Mitigation in O-RAN's RAN Intelligent Controllers', 2023, doi: 10.48550/ARXIV.2311.17482.
[169] C. Adamczyk and A. Kliks, 'Conflict Mitigation Framework and Conflict Detection in O-RAN Near-RT RIC', *IEEE Communications Magazine*, vol. 61, no. 12, pp. 199–205, Dec. 2023, doi: 10.1109/MCOM.018.2200752.
[170] O-RAN, 'O-RAN OTICS'. Accessed: Aug. 26, 2025. [Online]. Available: https://www.o-ran.org/otics
[171] O-RAN, 'WG3 TR "Near-Real-time RAN Intelligent Controller and E2 Interface: Conflict Mitigation"R004 v01.00', Alfter, Germany, O-RAN.WG3.TR.ConMit-R004-v01.00, Oct. 2024.
[172] M. Corici, R. Modroiu, F. Eichhorn, E. Troudt, and T. Magedanz, 'Towards efficient conflict mitigation in the converged 6G Open RAN control plane', *Ann. Telecommun.*, vol. 79, no. 9, pp. 621–631, Oct. 2024, doi: 10.1007/s12243-024-01036-2.
[173] N. R. De Oliveira, I. M. Moraes, D. S. V. De Medeiros, M. A. Lopez, and D. M. F. Mattos, 'An Agile Conflict-Solving Framework for Intent-Based Management of Service Level Agreement', in *2023 2nd International Conference on 6G Networking (6GNet)*, Paris, France: IEEE, Oct. 2023, pp. 1–8. doi: 10.1109/6GNet58894.2023.10317714.
[174] A. Zolghadr, J. F. Santos, L. A. DaSilva, and J. Kibiłda, 'Learning and Reconstructing Conflicts in O-RAN: A Graph Neural Network Approach', 2024, *arXiv*. doi: 10.48550/ARXIV.2412.14119.
[175] P. B. del Prever *et al.*, 'PACIFISTA: Conflict Evaluation and Management in Open RAN', 2024, *arXiv*. doi: 10.48550/ARXIV.2405.04395.
[176] A. E. Giannopoulos, S. T. Spantideas, G. Levis, A. S. Kalafatelis, and P. Trakadas, 'COMIX: Generalized Conflict Management in O-RAN xApps—Architecture, Workflow, and a Power Control Case', *IEEE Access*, vol. 13, pp. 116684–116700, 2025, doi: 10.1109/ACCESS.2025.3585774.
[177] VERGE, 'D1.1 Use cases, requirements and initial system architecture', Jul. 2023. [Online]. Available: https://drive.google.com/file/d/1AsIv1ir0R86Y0IKW1EdNsbkr0VLV85fU/view
[178] ACROSS, 'D2.1 Ecosystem definition, requirements and use cases', Dec. 2023. [Online]. Available: https://across-he.eu/wp-content/uploads/2024/10/ACROSS-D2.1-Ecosystem-definition-requirements-and-use-cases.pdf
[179] 6G-INTENSE, 'D2.1 Report on Requirements and PoC scenarios definition.', Oct. 2024. [Online]. Available: https://6g-intense.eu/wp-content/uploads/2024/12/6G-INTENSE_WP2_D2.1.pdf
[180] ETHER, 'D2.4 Final report on ETHER network architecture, interfaces, and architecture evaluation.', Aug. 2024. [Online]. Available: https://www.ether-project.eu/wp-content/uploads/sites/100/2024/10/ETHER_Deliverable_D2.4_V1.0_final.pdf

Consortium

6G LEADER

LEARNING-DRIVEN AND EVOLVED RADIO FOR 6G COMMUNICATION SYSTEMS

Funded by the European Union

6GSNS

Funded by the European Union. The project is supported by Smart Networks and Services Joint Undertaking (SNS JU) and its members. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or SNS JU. Neither the European Union nor the granting authority can be held responsible for them.