



Smart, Automated, and Reliable Security Service Platform for 6G

Deliverable D2.2

Use Cases, Requirements, ROBUST-6G Initial Architecture and Initial ROBUST-6G Dataspace



ROBUST-6G project has received funding from the [Smart Networks and Services Joint Undertaking \(SNS JU\)](#) under the European Union's [Horizon Europe research and innovation programme](#) under Grant Agreement No 101139068.

Date of delivery: 20/12/2024
Project reference: 101139068
Start date of project: 01/01/2024

Version: 1.0
Call: HORIZON-JU-SNS-2023
Duration: 30 months



Document properties:

Document Number:	D2.2
Document Title:	Use Cases, Requirements, ROBUST-6G Initial Architecture and Initial ROBUST-6G Dataspace
Editor(s):	José María Jorquera Valero (UMU), Manuel Gil Pérez (UMU)
Authors:	Contributors and their organization are listed below
Contractual Date of Delivery:	31/12/2024
Dissemination level:	PU
Status:	Final
Version:	1.0
File Name:	ROBUST-6G D2.2_v1.0

Revision History

Revision	Date	Issued by	Description
0.1	19.02.2024	ROBUST-6G WP2	Initial draft with ToC
0.2	15.04.2024	ROBUST-6G WP2	First draft with the description of the UCs and their scenarios
0.21	03.05.2024	ROBUST-6G WP2	Internal revision and comments of UCs and scenarios
0.22	10.05.2024	ROBUST-6G WP2	Second draft with revised version of UCs and scenarios
0.3	17.05.2024	ROBUST-6G WP2	First round of requirements and ROBUST-6G dataspace
0.4	31.05.2024	ROBUST-6G WP2	First complete draft with revised comments on UCs and requirements
0.5	11.09.2024	ROBUST-6G WP2	First high-level architecture proposal and ToC proposal
0.6	30.11.2024	ROBUST-6G WP2	First complete draft
0.7	06.12.2024	ROBUST-6G WP2	Second draft with a new further revised version
0.6	16.12.2024	ROBUST-6G WP2	Final complete draft after internal review
1.0	20.12.2024	ROBUST-6G WP2	Final version

Abstract

This deliverable features the complete definition of the three use cases that will boost the decisions and developments of the technical drivers of ROBUST-6G concerning data management and governance, trustworthy and sustainable Artificial Intelligence (AI) techniques, zero-touch security management, and physical layer security. They will shape the proposed objectives for achieving robust security and trustworthiness in sixth generation (6G) networks. Along with the use cases and project objectives, a series of requirements are set out to serve as the basis for the design and development of ROBUST-6G components. Their fulfilment aims to build and present an initial architecture that facilitates seamless collaboration across multiple network domains while ensuring compliance with privacy regulations and tackling the complexities of cyber threats. Finally, a dataspace architecture is also detailed in this deliverable, which serves as a foundational element for the development of the project components to ensure that security and trustworthiness are embedded throughout the network infrastructure.

Keywords

Use cases, trustworthiness, requirements, ROBUST-6G architecture, distributed AI-driven security, physical layer security, zero-touch security management, data management and governance

Disclaimer

Funded by the European Union. The views and opinions expressed are however those of the author(s) only and do not necessarily reflect the views of ROBUST-6G Consortium nor those of the European Union or Horizon Europe SNS JU. Neither the European Union nor the granting authority can be held responsible for them.

List of Contributors

Participant	Short Name	Contributors
Ericsson Arařtırma Geliřtirme ve Biliřim Hizmetleri A.ř	EBY	Güneř Kesik, Ahmet Cihat Baktır, Ömer Faruk Tuna, Leyli Karaçay, Betül Güvenç Paltun, Hakan Alakoca, Bilal Çiçek, řamil Karaman
Telefónica Innovación Digital	TID	Lucía Cabanillas Rodríguez, Ignacio Domínguez, Diego R. López
Universidad de Murcia	UMU	Pedro M. Sánchez Sánchez, Enrique Tomás Martínez Beltrán, Alberto García Pérez, José María Jorquera Valero, Manuel Gil Pérez
Chalmers University of Technology	CHA	Tommy Svensson
University College Dublin	UCD	Bartłomiej Siniarski, Chamara Sandeepa, Thulitha Senevirathna, Farah Abed Zadeh
University of Padova	UNIPD	Stefano Tomasin, Giovanni Perin
Nextworks	NXW	Enrico Alberti, Pietro Giuseppe Giardina, Marco Ruta
ENSEA/Cergy	ENSEA/CERGY	Sara Berri, Arsenia Chorti
EURECOM	EUR	Ioannis Pitsiorlas, Marios Kountouris
GOHM Elektronik ve Biliřim San. Tic. Ltd. řti.	GOHM	Cem Ayyıldız, Fatih Emre Yıldız, Veli Can Yıldırım

List of Reviewers

Participant	Short Name	Contributors
Ericsson Arařtırma Geliřtirme ve Biliřim Hizmetleri A.ř	EBY	Betül Güvenç Paltun, Mustafa Riza Akdeniz
GOHM Elektronik ve Biliřim San. Tic. Ltd. řti.	GOHM	Cem Ayyıldız

Executive Summary

This deliverable describes the main findings and identification of the ROBUST-6G use cases aimed at ensuring the security and trustworthiness of the sixth generation (6G) network through the development of (i) integrated Artificial Intelligence (AI) / Machine Learning (ML)-driven solutions; (ii) a zero-touch security management solution to tackle multiple cyber-physical threats in Internet of Things (IoT) environments; (iii) a solution to guarantee physical layer security through AI/ML-driven technologies; and a (iv) Data Management Platform to enable and oversee the entire flow of data within the ROBUST-6G dataspace.

The three use cases and associated scenarios for ROBUST-6G are the following:

1. Use Case 1 “*AI model trustworthiness evaluation for 6G distributed scenarios*”, which emphasises the evaluation of AI/ML models in distributed 6G networks using Decentralized Federated Learning (DFL). It addresses key dimensions such as robustness, sustainability, explainability, and fairness while integrating physical layer security measures.
2. Use Case 2 “*Automatic threat detection and mitigation in 6G-enabled IoT environments*”, primarily focuses on the automation of security management processes within 6G networks. It aims to develop mechanisms that enable real-time threat detection and response, leveraging the programmability and flexibility of 6G to enhance security orchestration.
3. Use Case 3 “*Security capabilities exposure with Network-Security-as-a-Service (NetSecaaS)*”, which addresses the challenges of data governance in 6G environments, emphasising the need for effective data management strategies that ensure compliance with privacy regulations.

Following a detailed description of the proposed use cases, and tightly aligned with them, a significant list of requirements that the subsequent initial architecture must comply with is enumerated: both functional and non-functional requirements; technical requirements containing details about the technology stack or infrastructure; operational and business requirements about features that must be fulfilled from the point of view of the system stakeholders; and end user (customer) requirements of ROBUST-6G. This list of requirements is structured in different application domains with the aim of covering the main developments proposed in the project: data management and governance, trustworthy and sustainable AI techniques, zero-touch security management, and physical layer security. Also, an initial number of global requirements with cross-cutting conditions to be fulfilled are also described.

The above is the entry point for the building and design of an initial architecture for the ROBUST-6G project, which meets the demanded requirements as well as considering the different use cases and scenarios proposed for its subsequent implementation. This initial architecture is designed to support the three use cases defined by providing a modular and flexible framework that integrates various technologies, including AI/ML-driven solutions, continuous monitoring, and effective data management capabilities, among others. In addition, this deliverable also introduces the concept of dataspace as a vital framework for data management in 6G networks to ensure data security, compliance, and integrity, addressing challenges in managing sensitive information. The associated modules making up the dataspace allow creation of a secure and efficient environment for data management, supporting privacy-preserving solutions across multiple domains.

Table of Contents

1	Introduction.....	12
1.1	Motivation, objectives, and scope.....	12
1.2	Document structure.....	13
2	Use cases.....	13
2.1	AI model trustworthiness evaluation for 6G distributed scenarios.....	13
2.1.1	Motivation and overall description.....	13
2.1.2	State-of-the-art for application in 6G networks.....	15
2.1.3	Use case detailed description.....	15
2.1.4	Scenarios.....	17
2.2	Automatic threat detection and mitigation in 6G-enabled IoT environments.....	19
2.2.1	Motivation and overall description.....	19
2.2.2	Stakeholders definition, roles, and interactions.....	21
2.2.3	State-of-the-art for application in 6G networks.....	22
2.2.4	Use case detailed description.....	22
2.2.5	Scenarios.....	23
2.3	Security capabilities exposure with Network-Security-as-a-Service (NetSecaaS).....	26
2.3.1	Motivation and overall description.....	26
2.3.2	Stakeholders definition, roles, and interactions.....	27
2.3.3	State-of-the-art for application in 6G networks.....	28
2.3.4	Use case detailed description.....	30
2.3.5	Main scenario description.....	31
3	ROBUST-6G requirements.....	32
4	ROBUST-6G architecture.....	40
4.1	High-level ROBUST-6G architecture.....	40
4.2	Functional architecture of ROBUST-6G.....	41
4.3	High-level deployment view of ROBUST-6G.....	43
4.4	ROBUST-6G security services in the architecture.....	44
4.4.1	Data Management Platform.....	44
4.4.2	Trustworthy and Sustainable AI Services.....	45
4.4.3	Zero-touch Security Management.....	51
4.4.4	Physical Layer Security.....	53
4.4.5	Use Case Interactions.....	54
5	ROBUST-6G dataspace.....	58
5.1	Data fabric.....	59
5.2	Data governance.....	60
5.2.1	Data Catalog.....	60
5.2.2	Data Security.....	62
6	Conclusion.....	66

List of Tables

Table 3-1: Structure of the tables with requirements.....	32
Table 3-2: Overall system requirements for the ROBUST-6G system	33
Table 3-3: Requirements of the physical layer security in the ROBUST-6G system	34
Table 3-4: Data management requirements in the ROBUST-6G system.....	35
Table 3-5: Distributed AI-driven security requirements in the ROBUST-6G system	37
Table 3-6: Zero-touch security management requirements in the ROBUST-6G system	38

List of Figures

Figure 2-1: Use Case 1 flow diagram.....	17
Figure 2-2: AI model trustworthiness evaluation diagram for 6G distributed scenarios.....	18
Figure 2-3: ROBUST-6G components interacting with the external world.....	21
Figure 2-4: Device violation to cause an economic harm (a).....	24
Figure 2-5: Fraudulent usage of device resources	25
Figure 2-6: Device violation to cause an economic harm (b).....	26
Figure 2-7: Integration of ROBUST-6G with Open Gateway.....	27
Figure 2-8: NaaS ecosystem, roles and usage of APIs	29
Figure 2-9: Use Case 3 system architecture	31
Figure 4-1: High-level architecture of ROBUST-6G project.....	41
Figure 4-2: Functional architecture of ROBUST-6G project.....	42
Figure 4-3: High-level deployment view of ROBUST-6G project contributions	43
Figure 4-4: Data Fabric architecture.....	45
Figure 4-5: ROBUST-6G Zero-touch security platform functional architecture	52
Figure 4-6: UC1 Scenario 1 interactions	55
Figure 4-7: UC2 high-level functionalities interactions	56
Figure 4-8: UC3-1 interactions.....	57
Figure 4-9: UC3-2 interactions.....	57
Figure 4-10: UC3-3 interactions.....	58
Figure 5-1: Dataspace architecture	58
Figure 5-2: Data modelling	59
Figure 5-3: Data Fabric	59
Figure 5-4: Conceptual metamodel of the Data Catalog	60
Figure 5-5: Combination of DCAT and RDF datasets (source: [Ate16]).....	61
Figure 5-6: Access control mechanism	63
Figure 5-7: Data flow for access control mechanism	64
Figure 5-8: Data Security API.....	65
Figure 5-9: Provenance.....	66

Acronyms and abbreviations

Term	Description
3GPP	3rd Generation Partnership Project
5G	Fifth Generation
6G	Sixth Generation
AAA	Authentication, Authorization, and Accounting
AI	Artificial Intelligence
AIaaS	AI-as-a-Service
AL	Accounting Ledger
AoA	Angle of Arrival
AoI	Age of Information
API	Application Programming Interface
ASP	Application Service Provider
B5G	Beyond Fifth Generation
BS	Base Station
BSS	Business Support System
CBOR	Concise Binary Object Representation
CFL	Centralized Federated Learning
CI/CD	Continuous Integration and Continuous Delivery/Deployment
CLEVER	Cross-Lipschitz Extreme Value for nEtnetwork Robustness
CN	Core Network
CNN	Convolutional Neural Network
COSE	CBOR Object Signing and Encryption
CPU	Central Processing Unit
CSI	Channel State Information
DCAT	Data Catalog Vocabulary
DDoS	Distributed Denial of Service
DFL	Decentralized Federated Learning
DL	Deep Learning
dMIMO	distributed Multi-Input Multi-Output
DoA	Description of Action
DoS	Denial of Service
DP	Differential Privacy
E2E	End-to-End
EDPB	European Data Protection Board

ETL	Extract-Transform-Load
ETSI	European Telecommunications Standards Institute
FAIR	Findability, Accessibility, Interoperability, and Reusability
FedAvg	Federated Averaging
FL	Federated Learning
FOAF	Friend-of-a-Friend Ontology
GDPR	General Data Protection Regulation
GPU	Graphics Processing Unit
GSMA	Global System for Mobile Association
GUI	Graphical User Interface
HE	Homomorphic Encryption
ICT	Information and Communication Technologies
IdP	Identity Provider
IDS	Intrusion Detection System
IETF	Internet Engineering Task Force
IoT	Internet of Things
IPsec	Internet Protocol security
ISP	Internet Service Provider
ISV	Independent Software Vendor
JSON	JavaScript Object Notation
KGC	Knowledge Graph Construction
KPI	Key Performance Indicator
KVI	Key Value Indicator
LCM	Lifecycle Management
LDAP	Lightweight Directory Access Protocol
LIME	Local Interpretable Model-Agnostic Explanations
LOT	Linked Open Terms
LSTM	Long Short-Term Memory
MIMO	Multi-Input Multi-Output
MITM	Man-in-the-Middle
ML	Machine Learning
mMIMO	Massive Multi-Input Multi-Output
MNO	Mobile Network Operator
MQTT	Message Queuing Telemetry Transport
MWC	Mobile World Congress
NaaS	Network-as-a-Service

NetSecaaS	Network-Security-as-a-Service
NF	Network Function
NSMgmt	Network Security Management
OAuth	Open Authorization
ODA	Open Digital Architecture
OGWTS	Open Gateway Technical Stream
OPAG	Operator Platform API Group
OPG	Operator Platform Group
ORG	Organization Ontology
OSS	Operational Support System
OT	Operational Technology
OTP	One-Time Password
OTT	Over-The-Top
PaC	Policy-as-Code
PAP	Policy Administration Point
PDP	Policy Decision Point
PEP	Policy Enforcement Point
PHY	Physical
PKI	Public Key Infrastructure
PLMN	Public Land Mobile Network
PLS	Physical Layer Security
PMP	Programmable Monitoring Platform
PNI-NPN	Public Network Integrated Non-Public Network
PoC	Proof of Concept
PROV-O	PROV Ontology
QoD	Quality on Demand
QoS	Quality of Service
RAN	Radio Access Network
RDBMS	Relational Database Management System
RDF	Resource Description Framework
RF	Radio Frequency
RFC	Request for Comments
RIS	Reflective Intelligent Surface
RML	RDF Mapping Language
RNN	Recurrent Neural Network
S-Application	Security Application

S-CL	Security Closed-Loop
S-CLMgmt	Security Closed-Loop Management
SDN	Software-Defined Networking
SecaaS	Security-as-a-Service
SHAP	Shapley Additive Explanations
SKA	Secret Key Agreement
SKG	Secret Key Generation
SKOS	Simple Knowledge Organization System
SMC	Secure Multiparty Computation
SME	Small and Medium Enterprise
SMPC	Secure Multi-Party Computation
SPARQL	SPARQL Protocol and RDF Query Language
SPARQL-SD	SPARQL Service Description
S-RO	Security Resource Orchestration
S-Services	Security Services
SSLA	Security Service Level Agreement
S-SO	Security Service Orchestration
UC	Use Case
UE	User Equipment
VAE	Variational Autoencoder
W3C	World Wide Web Consortium
WAS	Wholesale Agreement Services
WIMSE	Workload Identity in Multi-Service Environments
WP	Work Package
XAI	Explainable Artificial Intelligence
ZSM	Zero-touch Service Management
ZTS	Zero-Touch Security

1 Introduction

The advent of sixth generation (6G) networks entails a pivotal step in the evolution of wireless communication, promising cutting-edge capabilities in connectivity, data processing, intelligent automation, security, and trustworthy Artificial Intelligence (AI) models. As global industries and critical systems increasingly rely on advanced digital infrastructures, 6G is poised to address emerging challenges that extend beyond conventional Key Performance Indicators (KPIs) such as high-speed connectivity, bandwidth, and throughput. In particular, novel Key Value Indicators (KVI) [URB+21] are gaining prominence in forthcoming architecture designs so as to lay the foundation of 6G technologies. Specially, trustworthiness, resilience, and security are emerging as foundational pillars in the design and development of 6G-oriented solutions, shaping the future of applications across multiple administrative domains.

To begin with, the trustworthiness of AI models in distributed environments such as cross-domain or Cloud Continuum scenarios presents a groundbreaking research area for 6G contributions. The use of Decentralized Federated Learning (DFL) techniques may offer a promising pathway for elaborating joint privacy-preserving AI models that can be trained in a distributed way. In this way, the need to address data privacy and distributed intelligence deployment is a dual demand to be tackled [KMA+21]. Furthermore, ensuring trustworthiness at the physical and sensing layer is another imperative challenge to secure the underlying infrastructure of 6G networks, necessitating advanced mechanisms to safeguard against adversarial cyber threats and maintain data integrity [MJC+21].

Second, as the proliferation of 6G-enabled Internet of Things (IoT) systems amplifies the complexity of network environments, the need for robust threat detection and mitigation strategies becomes paramount. In this vein, security orchestrators, acting as centralized or decentralized control systems, can leverage 6G's advanced capabilities to enable real-time responses to cyber threats, thereby enhancing the resilience of IoT ecosystems [GYZ+21]. Likewise, new opportunities may emerge in the literature since 6G networks require the capacity of on-demand deciding the assets to be secured which may be deployed in several administrative domains or network segments. In this regard, the flexibility and programmability characteristics play a fundamental role. It will enable management and orchestration components to determine the most appropriate monitoring tools to fulfil the security requirements declared by users and find out potential threats happening in real time. On another hand, risk-averse resource management solutions may help security orchestrators to minimise the threats and privacy leaks, as embedding algorithms that evaluate potential risks could be introduced into orchestration activities.

Finally, 6G networks also offer a unique opportunity to expose security capabilities through innovative frameworks such as Network-Security-as-a-Service (NetSecaaS). By integrating security functionalities directly into the network infrastructure [OD22], NetSecaaS may empower organisations to tailor their security strategies to meet dynamic threats, fostering a proactive approach to cyber defence. However, improving standardised Application Programming Interfaces (APIs) are developed to abstract security capabilities, thereby enabling application developers and enterprises to apply security policies without requiring deep network expertise.

1.1 Motivation, objectives, and scope

ROBUST-6G aims to contribute to the design and development of reliable and security AI-driven solutions for 6G networks by addressing critical challenges such as trustworthiness, security, and scalability. To achieve this aim, ROBUST-6G will analyse and integrate advanced technologies like DFL, explainable AI, zero-touch security management, continuous monitoring, threat detection and prediction, physical layer security, and data governance.

Therefore, the objectives and scope of this deliverable are to pave the way with respect to forthcoming design and development activities to be carried out in the next technical Work Packages (WPs). One of the objectives of this document is to clearly determine what are the main use cases and scenarios that will prove ROBUST-6G designs and deployments. For this reason, Deliverable 2.2 introduces three different use cases in which the main topics of the ROBUST-6G project are consequently covered: data management and governance, trustworthy and sustainable AI techniques, zero-touch security management, and physical layer security. Additionally, each use case proposes two or more scenarios in which the component and micro-services defined in the ROBUST-6G architecture can be tested. Linked to the UCs and the associated technical WPs, five domains (Global, Physical Layer, Distributed AI-driven Security, and Zero-touch Security Management)

group the agreed requirements that will be the basis for the design and development of our ROBUST-6G components.

On another hand, considering all the previous information, high-level and functional architecture views are presented in this deliverable. The main objective is to offer a modular perspective on the technologies planned for development within the project, highlighting the system's goal of achieving end-to-end, integrated security for the envisioned 6G networks. As well, the aforementioned main topics display further high-level and functional details to reflect how they go deeper into the initial design declared in the ROBUST-6G architecture while maintaining the principles and interactions already settled. Last but not least, objective of this deliverable is to contextualise the dataspace which is in turn divided into two modules: Data Fabric and Data Governance. ROBUST-6G dataspace boosts the power and role of intelligent data management in a distributed environment where several data sources may feed different types of data that can be correlated to obtain more for more sophisticated querying and data analysis via knowledge graphs and appropriate control of the data products.

1.2 Document structure

The document at hand is structured as follows. In Section 2, three main use cases, their scenarios, and stakeholders are thoroughly described to point out some of the principal pillars of the ROBUST-6G project, trustworthy AI models for decentralized environments, e.g., physical and sensing layers, automatic threat detection and mitigation in 6G IoT environments, and Network-Security-as-a-Service. Section 3 collects a set of (non-)functional, operational, and business requirements for five application domains. Afterward, Section 4 presents ROBUST-6G architecture, high-level, functional, and deployment views, in which key modules, components, and services of our work packages are delineated. In addition, the utmost important interactions and communications between modules, components, and services can also be observed. Section 5 describes ROBUST-6G dataspace in charge of managing data flows to build a secure and efficient environment for data management across distributed domains. Finally, we summarise our conclusions in Section 6.

2 Use cases

This section offers an initial, but comprehensive presentation of the three Use Cases (UCs) that we initially defined in ROBUST-6G's Description of Action (DoA), incorporating several application scenarios in one of them. For each UC and its potential scenarios, a general motivation and a high-level description of its main objectives, the list of stakeholders involved and their interaction with different high-level workflows, where possible, as well as a first set of expected KPIs and KVIs for each scenario are provided. It should be noted that the selection of the UCs was made taking into consideration the complementarity in the different technical aspects proposed in ROBUST-6G.

2.1 AI model trustworthiness evaluation for 6G distributed scenarios

The first UC presented below focuses on DFL for training AI and Machine Learning (ML) models, incorporating trust dimensions such as robustness, sustainability, explainability, and fairness, in addition to integrating Physical Layer Security (PLS) measures. Key results include AI/ML models that are robust to cyber threats, privacy-friendly and adaptable to dynamic 6G environments.

2.1.1 Motivation and overall description

The decentralized nature of forthcoming 6G networks presents cutting-edge challenges concerning the generation of shared AI/ML models while preserving privacy and fostering trust. These challenges arise from balancing the collaboration required for model training with the stringent demands for privacy and security. The essence of this problem lies in establishing robust mechanisms for trustworthy and efficient decentralized learning, which are pivotal to unlocking the potential of 6G. In this context, a 6G network or domain can be delineated as the collective networks falling within the administrative purview of a certain entity, encompassing all network nodes within its infrastructure. This definition underscores the complex, interconnected environment in which decentralized AI/ML model generation must operate, requiring sophisticated approaches to collaboration and security. Thus, the development of AI/ML models necessitates a paradigm that supports collaboration across diverse nodes within each domain while upholding stringent privacy standards. DFL) enables collaboration between different nodes by exchanging model updates rather than raw data, enabling the training of shared models in a secure and distributed manner. The generation of AI/ML models entails this collaborative effort, leveraging DFL to balance the need for privacy with the

efficiency of shared learning. Throughout the training process and upon the completion of model training, the trustworthiness of the models must be meticulously evaluated before their deployment on end devices for production purposes. This first UC of the ROBUST-6G project is focused on evaluating the performance and reliability of these shared and decentralized AI/ML models. This evaluation covers three critical dimensions: (i) the trustworthiness of the AI/ML models based on essential pillars, primarily model *robustness*, *sustainability*, *explainability*, and *fairness*; (ii) the environment and context in which the AI/ML models were generated, such as reputation relations between networks and domains, the use of secure channels for communications; and (iii) trustworthiness measures of the infrastructure layer, i.e., the physical and sensing layers, together with the proposed PLS-based mitigation of potential attacks. The first two aspects are clearly covered below in the first scenario of Section 2.1.4.1, while the third one will be addressed in more detail in the second scenario described in Section 2.1.4.2. Stakeholders definition, roles, and interactions. One of the keys to any UC is to identify and define the stakeholders involved, such as domain administrators and end users, and outline their roles—from contributing data to models, to evaluating the trustworthiness of these models. Furthermore, the document delineates the interactions between different nodes (cloud, edge, and extreme edge) which may occur horizontally or vertically, reflecting the decentralized nature of the model training process. In the context of developing trustworthy AI for DFL in 6G networks, it is crucial to clearly define the various stakeholders involved. These stakeholders play pivotal roles in ensuring that the AI systems are effective and adhere to ethics and security standards. Below, the key stakeholders in this UC are identified, each with specific responsibilities and expectations that contribute significantly to the overall success and trustworthiness of the AI models developed. Individuals or entities who manage a domain. A domain encompasses all network nodes under the administrative scope of an entity. Administrators are responsible for overseeing the AI/ML model training and trustworthiness evaluations within their respective domains. This stakeholder is represented by Telecom Operators, Internet Service Providers (ISPs), Data Center Administrators, etc.

- **AI Developers.** They design and develop the algorithms for DFL models. They are tasked with integrating AI models into existing network structures and ensuring they perform optimally across different nodes. This stakeholder is represented by AI labs and AI research teams involved in the telecom network environments.
- **End Users.** They are the primary beneficiaries of DFL in 6G networks. They rely on trustworthy AI models for decision-making and benefit from personalized services and enhanced data privacy. These end users can be individual or enterprise-level customers.
- **Federation Devices.** These represent the operational stakeholders within the DFL ecosystem. These devices, including cloud, fog, edge, and extreme edge nodes, such as mobile base stations, cloud servers, smartphones, and other end-user devices or infrastructure computing systems, actively participate in training and sharing AI/ML model updates. Each device plays a crucial role in enabling the distributed learning process, contributing to developing and refining the final models while maintaining the decentralized architecture.
- **Regulatory bodies.** These entities enforce standards and regulations that govern the ethical use of AI, data privacy, and security in Federated Learning (FL) environments, such as the European Data Protection Board (EDPB) that oversees compliance with the General Data Protection Regulation (GDPR), which directly impacts AI applications that process personal data, or the European Telecommunications Standards Institute (ETSI) which establishes standards for AI integration within telecom networks.

Next, the classification of nodes is based on their roles within the framework, which defines the set of responsibilities and actions assigned to individuals or groups in the ecosystem. Each role is critical to managing the complex interactions and processes for developing and maintaining AI systems.

- **Model Contributor.** Nodes (cloud, fog, edge, extreme edge) that participate in generating and sharing updates for AI/ML models. They train local models on their data without exposing it, contributing to the privacy-preserving aspect of FL.
- **Trust Evaluator.** Both domain administrators and potentially automated systems that assess the AI/ML models on trust dimensions such as robustness, fairness, and explainability, using aggregated scores to determine overall trustworthiness.
- **Security and Privacy Monitor.** These roles are dedicated to ensuring that the communication channels between nodes are secure and that data privacy is maintained throughout the learning process. This role could be assumed by stakeholders representing Domain and Network Administrators.

Regarding interactions, they refer to the dynamic processes and communications between various stakeholders and system components. These interactions are crucial for the synchronisation and functionality of the learning process across distributed networks. They encompass everything from data sharing and model updates to the evaluation of trustworthiness and security measures. Understanding these interactions is key to optimising the AI system's performance and ensuring its reliability and integrity.

- **Model Sharing.** Can be *horizontal*, where devices of the same level (e.g., cloud to cloud) share model updates, or *vertical*, involving different levels (e.g., cloud to extreme edge). This structure supports the decentralized nature of the learning process.
- **Trustworthiness Evaluation.** Once the final models are trained, they are evaluated for trust dimensions. This includes not only the performance of the models but also the processes through which they were created and the reputation of the participating domains.
- **Reputation Assessment.** Trustworthiness evaluations also incorporate assessments of the reputation relationships between nodes and domains. This is crucial in environments where collaboration is necessary but challenging due to the autonomous nature of each domain and potential security concerns.

2.1.2 State-of-the-art for application in 6G networks

Current solutions for AI trustworthiness evaluation are centred on evaluating the model and its configuration once it has been trained in a centralized manner, where data coming from the different nodes are joined together for processing or where each node generates its own model using local data [ZLQ+20]. However, these solutions are unsuitable for modern 6G scenarios due to the highly distributed network topologies, the large number of nodes, and stringent privacy requirements. Furthermore, 6G is expected to incorporate decentralized federated schemes where central entities play a minimal role, aligning with the shift toward distributed AI model generation. Trustworthiness evaluation of AI/ML models generated under the DFL paradigm is still an open challenge due to the particularities in these setups, such as the lack of knowledge of the training data, malicious node presence or node participation ratio [BM21]. In addition, nowadays, there is no integration between the trust in AI model performance and the trust between the domains or entities participating in the generation of these models. This is a key challenge for trustworthy 6G networks, as these two trust dimensions must be merged to have a complete view of the trustworthiness of the models before they are elevated into production deployment.

Several solutions in the existing literature have already worked on frameworks for the trustworthiness evaluation of centralized AI/ML models. These frameworks calculate a score for the following pillars of trust: *accountability*, *fairness*, *explainability*, and *robustness* [WL24]. Typically, the evaluation process involves designated entities within the framework—such as trusted evaluation servers or model auditing tools—calculating individual scores for each pillar. These scores are then aggregated, often on a central evaluation node or server, to produce a final trustworthiness score for the model. As part of the development of the first scenario of this UC, detailed below, the expected framework will be extended to support new trust metrics and pillars arising from the usage of DFL to generate AI/ML models.

In the same perspective, other solutions in the literature worked on approaches to evaluate the reputation relationships between different domains in the Fifth Generation (5G) and Beyond Fifth Generation (B5G) networks [JSG+22]. In this regard, end users may consider historical data and monitoring data provided by trustworthy data repositories to assess the reputation of other domains. Reputation approaches should be capable of handling both an entity reputation with which the updated models are to be shared and the reputation of the domain per se. Therefore, the reputation of a given entity/node affects all overall domains to which it belongs. Furthermore, these approaches need to consider security and privacy aspects, for instance, the involvement of the node in adversary activities or the ability of the node to maintain proper security measures against external fraudulent activity. Thereby, current approaches will be updated and enhanced to integrate them into the AI model trustworthiness evaluation lifecycle, using the reputation of the domains as one of the inputs in the process.

2.1.3 Use case detailed description

This UC describes a scenario where multiple domains, each managing distinct network nodes (including cloud, edge, and extreme edge nodes), collaborate to develop shared AI/ML models. The objective is to harness DFL techniques to ensure the privacy, security, and trustworthiness of these AI models across highly distributed networks.

Domains operate autonomously but share the objective of training common AI models without centralising data. This approach preserves user privacy and data integrity. Each domain contributes to the model training by processing data locally at its nodes and sharing model updates with peers instead of raw data. This collaborative training involves both horizontal interactions (e.g., cloud-to-cloud) and vertical interactions involving different levels (e.g., cloud-to-edge). Key challenges include safeguarding data privacy, maintaining integrity, and ensuring model reliability and fairness despite limited direct data access. The distributed network nature also raises coordination and synchronisation complexities, increasing vulnerability to security threats such as data breaches or adversarial attacks on AI models.

The interactions in this UC are diverse. Nodes exchange updates to AI models, promoting collaborative yet private training. Models are evaluated on various metrics to determine deployment readiness. Security measures are continuously monitored and updated to protect data and model integrity.

The primary goal is to develop a framework where AI models are effective in their predictive capabilities and exemplary in their ethical, fair, and transparent use. This system embeds trust at every stage of the AI model lifecycle, from data collection to final deployment, ensuring models meet the complex needs of 6G networks.

Figure 2-1 outlines a comprehensive process flow for implementing AI model trustworthiness evaluation for 6G distributed scenarios. This flow includes detailed steps and stakeholder interactions that are essential for maintaining the integrity and trustworthiness of AI systems. The sequence of steps begins with the initiation of FL operations and covers the iterative cycles of model training, sharing, and refinement. It also incorporates crucial evaluations of model trustworthiness and regulatory compliance, culminating in the integration of user feedback and final model deployment. Each phase is designed to ensure collaborative engagement among all stakeholders, from network administrators to end users, thereby facilitating a robust framework for AI deployment in decentralized networks. This structured process not only addresses the technical complexities associated with 6G technologies but also ensures that the AI systems deployed are both effective and ethically sound. The steps identified in this process are:

1. Initialization of FL Process: Domain and Network Administrators launch the FL operations, coordinating with AI Developers to deploy algorithms across the network.
2. FL Cycle:
 - Local Model Training: Federation Nodes at the cloud, edge, and extreme edge independently train local models using their specific datasets.
 - Model Sharing: Nodes share their AI/ML model updates with others, both horizontally among similar level nodes (e.g., cloud to cloud) and vertically between different levels (e.g., cloud to extreme edge).
 - Reputation Assessment: Domain and Network Administrators evaluate the inter-node reputation based on data shared during the model updates, aiding in the detection of any potential issues related to node reliability.
 - Partial Trust Assessment: Trust Evaluators conduct intermediate assessments of the models on trust dimensions such as robustness and fairness using the performance data provided during the cycle. This helps in making ongoing adjustments before the final model aggregation.
 - (Aggregated) Model Refinement: Nodes collaboratively refine an aggregated model through collective learning, monitored by Network Administrators for integrity and efficiency.
3. Trustworthiness Evaluation: After the FL cycles are complete, Trust Evaluators assess the final aggregated models on comprehensive trust dimensions, including explainability, and provide final trustworthiness scores.
4. (Regulatory) Compliance and Certification: Regulatory Bodies review and certify the processes and models according to legal and ethical standards, ensuring compliance is maintained throughout the learning process.
5. Final Reputation and Trust Assessment: Domain and Network Administrators conduct a final assessment of the reputation and performance of the participating domains and nodes. This assessment, coupled with the trust evaluations, guides future FL initiatives and collaborations.
6. Feedback Integration and Deployment: End Users provide feedback on the models' effectiveness, which is integrated by Domain and Network Administrators before final model deployment across the network. Effectiveness can be measured in terms of performance, Quality of Service (QoS), trustworthiness, etc. This feedback may also be shared with Regulatory Bodies to inform and refine compliance standards based on real-world user experiences. If the results are not satisfactory, this step

can trigger the relaunch of the entire process, ensuring the models are optimized and aligned with both user needs and regulatory requirements.

Each step and cycle are critical for maintaining the integrity, effectiveness, and trustworthiness of AI systems in a decentralized 6G environment. This structured approach ensures that all stakeholders are continuously engaged, and that the system adapts to new information and challenges collaboratively.

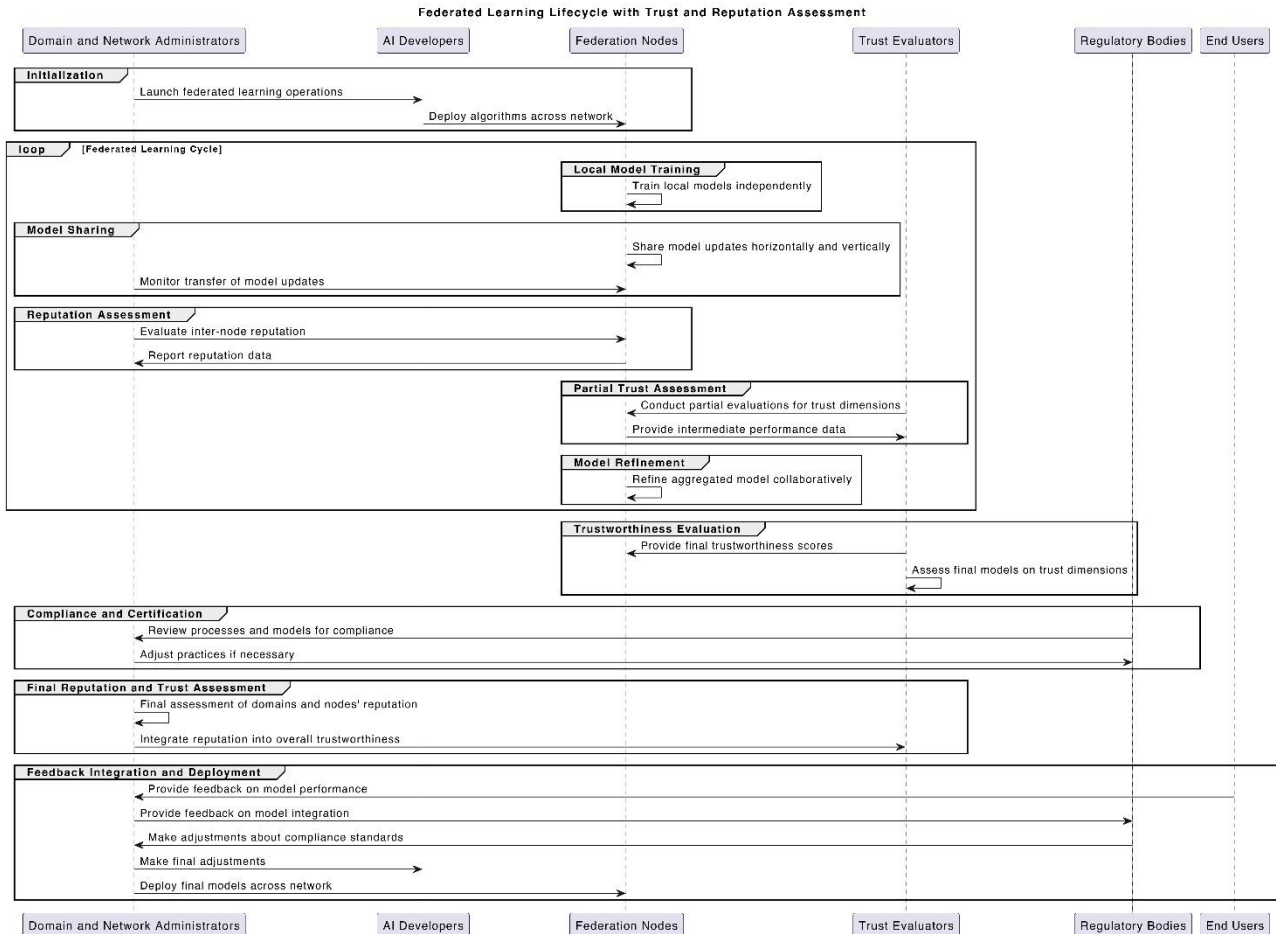


Figure 2-1: Use Case 1 flow diagram

Successful implementation will result in AI systems that are robust against threats, compliant with privacy standards, and efficient in the decentralized, dynamic 6G environment. This will enable new AI-driven applications in industries like healthcare, automotive, and public safety, offering more responsive, adaptive, and personalised solutions.

2.1.4 Scenarios

The two scenarios that make up this first UC are outlined below. The first scenario of Section 2.1.4.1 develops a DFL framework that emphasizes privacy, trustworthiness, and model robustness, taking advantage of the collaboration of the federation nodes. On the other hand, the second scenario of Section 2.1.4.2 focuses on reliability and resilience at the physical and sensing layers, using AI models to improve authentication, secret key agreements and threat detection through probabilistic measures and physical layer data integration.

2.1.4.1 Decentralized federated learning for joint privacy-preserving AI/ML model training

The main objective of this scenario is the design of a fully DFL framework to enable AI trustworthiness assessment in highly distributed network topologies. This framework steers the generation of AI models within federated schemes in which central entities such as servers are bypassed, as opposed to current centralized solutions that are not well suited to state-of-the-art distributed 6G scenarios. This scenario also seeks to assess the trustworthiness of AI/ML models following a DFL approach, deployed in the framework mentioned above, by analysing fundamental pillars such as accountability, fairness, explainability, and robustness, together with key aspects of FL such as privacy. In addition, this first scenario also aims to explore how the performance of

local AI/ML models can be improved if participants are able to weight, or even discriminate, model updates from other entities and networks based on past behaviour, i.e., pursuing a reputation-based trust approach.

Figure 2-2 showcases an example diagram for assessing the trustworthiness of AI models in a fully distributed and decentralized 6G scenario.

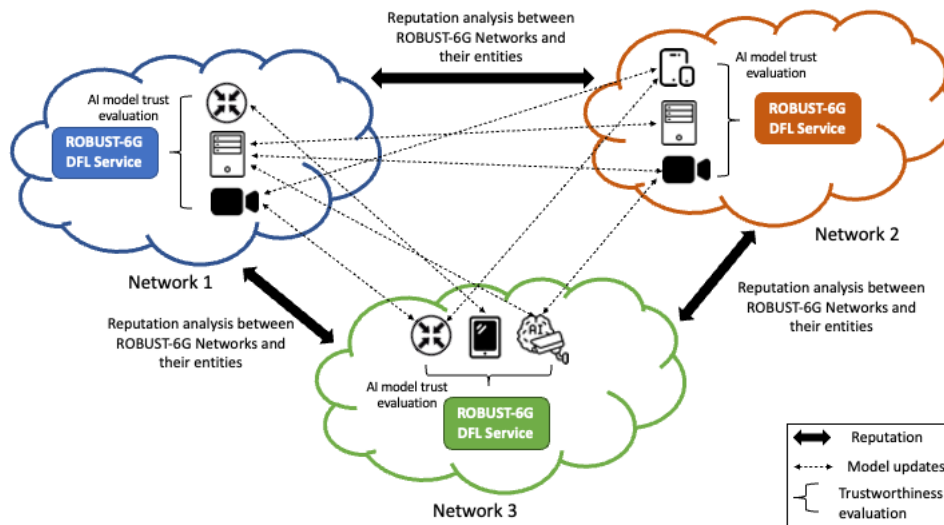


Figure 2-2: AI model trustworthiness evaluation diagram for 6G distributed scenarios

As depicted in Figure 2-2, different domains or networks collaborate to generate shared AI/ML models in a reliable and privacy-preserving manner. Consider that a network can be viewed as a given domain under the administrative scope of a certain entity, which encompasses all network nodes present in its infrastructure. Devices involved from different networks collaborate to generate numerous AI/ML models, and then the domain and network administrators compute the trust scores—different per domain—based on the model performance, the FL process and the reputation between the entities taking part in the process.

To generate the AI/ML models, the cloud, fog, edge, and extreme edge nodes of each network share the updates of the AI/ML models with the nodes of the other networks, training the shared models using a DFL approach. In each domain, nodes can directly interact with other domains or follow a hierarchical setup, where designated nodes act as proxies between domains. During the training process, and once the final model(s) has been trained, the AI/ML trustworthiness is assessed before deploying them on end devices.

Each network will then evaluate the final model, as well as the process through which it was generated. At this point, aspects of AI trustworthiness are considered: model robustness, explainability, fairness, accountability, and privacy, among others. In addition, other aspects related to the environment are examined in which the models were generated, such as reputation-related inter-network relationships as well as the use of secure communication channels for communications.

In this specific scenario, a given number of KPIs can be targeted to assess the performance and effectiveness of using a DFL approach, with the aim of fostering privacy preservation in AI/ML model building. Among these KPIs we can find:

- **Reliability:** Achieve a trustworthiness score of 80% or higher for each pillar (such as robustness, fairness, explainability, and accountability) assessing the holistic performance of DFL models. The score represents an aggregated evaluation across these pillars, with defined weights and calculation methods outlined earlier in the framework.
- **Model Accuracy:** Achieve an improvement in AI/ML accuracy of 5% or more compared with local training, on average, after considering the trust of entities and domains (networks) sharing their model updates.
- **DFL Robustness:** Guarantee that the DFL AI model has a minimum robustness score of 85% against adversarial attacks, measured in terms of Attack Success Rate or Cross-Lipschitz Extreme Value for nNetwork Robustness (CLEVER) score, which aim to subvert the proper operational functioning of the DFL framework.

2.1.4.2 *Physical and sensing layer trustworthiness and resilience*

This scenario aims at considering a new dimension provided by ROBUST-6G that incorporates trustworthiness measures from the infrastructure layer, specifically the physical and sensing layers. The information coming from the physical layer are, obtained from:

- Sensors on autonomous agents.
- Embedded Radio Frequency (RF) signatures in transmitted signals.
- Engineered RF fingerprints using distributed Multi-Input Multi-Output (dMIMO).
- Migrating RF Fingerprints between base stations.
- Other data, such as agents' positions.

From such measurements we aim at build mechanisms for security, in particular for authentication and secret key agreement. About authentication, by collecting several measurements from the environment and possibly controlling the environment itself (e.g., configuring Reflective Intelligent Surfaces–RIS) we aim at deciding if a message is coming from the legitimate transmitter or from an impersonating attacker. Such a decision is made by an artificial intelligent model that can also be shared among multiple users. The output is a score of authenticity which can be seen as a probabilistic measure. Also in this scenario, the interpretation of measures and their fusion with information coming from different layers will benefit from suitably trained AI models.

About secret key authentication we aim instead at obtaining a secret key among devices where the randomness of the key comes from the measurement of the shared channel at the physical layer. Also in this case, to obtain the secret key, AI models that have been properly trained are exploited at the devices. The transmission over the physical channel can also be controlled with beamformers in Multi-Input Multi-Output (MIMO) systems and RISs. Such controls can be used to improve the privacy and the overall trustworthiness of 6G networks.

This scenario will exploit information coming from the physical layer as well as the partial control of the physical layer itself to improve the security and trustfulness of 6G networks. The following KPIs highlight the key objectives for measuring reliability and resilience in the physical and sensing layers:

- Obtain an accuracy of the authentication mechanism of more than 90% in typical cellular scenarios.
- Obtain a key agreement rate of more than 99% before reconciliation mechanisms are used in Secret Key Agreement (SKA), in typical cellular scenarios.
- Model Accuracy: Achieve an improvement in AI/ML accuracy of 5% or more compared with local training for threat detection using PLS input.
- Accuracy: Achieve an 80% accuracy in detection of threats using RF fingerprinting solutions.

2.2 Automatic threat detection and mitigation in 6G-enabled IoT environments

This second UC explores threat detection and mitigation in 6G-enabled IoT environments, focusing on three specific scenarios inherent to small and medium-sized office environments, smart buildings and smart agriculture, respectively. The goal is to demonstrate how advanced closed-loop security mechanisms and AI-driven processes can address proactive, reactive, and predictive security needs, ensuring a resilient IoT ecosystem.

2.2.1 Motivation and overall description

In modern 6G networks, the increase in device connectivity generates a consistent rise in network traffic, which results in a new challenge to consider. To report some examples of these challenges, it is enough to think of common activities such as bandwidth usage optimization, latency-sensitive services management or QoS maintenance across applications like IoT. Lastly, yet significantly, maintaining security as the first line of defence and implementing protections is crucial since, in addition to legitimate requests, a significant percentage of fraudulent requests with numerous threats come together. Among the most famous cyber threats, it is worth mentioning Distributed Denial of Service (DDoS), cryptojacking, data breaches and unauthorized access attempts. On the other side, protection mechanisms such as AI-driven threat detection, encryption protocols and network isolation (slicing).

Concurrently, IoT is a fast-moving frontier. Numerous sensors are ready to help employees and companies with a wide range of everyday duties. However, the adoption of such utilities needs to consider the potential

hazards they generate as well. In these regards, malicious actors may try to exploit vulnerabilities in IoT infrastructures, which could result in large financial losses but also physical safety risks, operational disruptions, and compromised system reliability.

The following scenarios, described in detail in Section 2.2.4, focus on threat detection and mitigation in 6G-enabled IoT environments. These environments, as the name suggests, contain numerous sensors publishing data measurements into a shared network. The interconnection of a large amount of data, generated by combining network and sensor data information with a certain logic, facilitates the disclosure of new valuable pieces of information through a well-defined process known as data fusion and analysis. This process begins with the data collection, continues with preprocessing, integration, and analysis and concludes with the generation of new valuable insights. In industrial environments, companies adopt this strategy to gain new knowledge for their production process aiming at reducing their operational costs, increasing efficiency and improving the overall system performance. One standardized method on how to approach these improvements is through the concept of the closed-loop.

The Closed-Loop approach is a fundamental method that facilitates the identification and resolution of several problems. It works as a continuous cycle of the following four functions:

1. **Observation:** Data from networks and sensors are constantly checked. Network data such as system logs, sensor logs, user activities are fundamental for understanding if the target environment is working as expected. At the same time, sensor measurements provide additional valuable information about the environmental factors such as temperature, humidity, noise, and light.
2. **Analysis:** Data, especially in large volumes, does not provide value on its own. A deeper analysis of data (time, source, destination, actual meaning) is a difficult but necessary task. This function aims to discover possible threats or troubles in comparison to the ideal system lifecycle behaviour. This analysis may be executed by static processes, or as suggested by modern approaches, with the help of AI/ML solutions.
3. **Decision:** After the analysis identifies a threat, the decision step defines the most suitable response plan for addressing the problem. Decisions are taken based on predefined policies. The policies, or rules, aim at selecting the best actions to mitigate the identified problem. For example, if suspicious traffic is detected, then the decision step may suggest blocking the source following the predefined static rule. Furthermore, in more advanced systems, dynamic rules could be generated through AI-based analysis (e.g., adaptive response to emerging threats). Finally, it is important to underline once more the importance of the connection between the analysis and decision steps. The second, in fact, strictly relies on the intelligence actions derived from the analysis, which allows triggering the most effective policy to mitigate the identified threat.
4. **Action:** Finally, the chain of the closed loop terminates with the execution of an action aimed at solving the discovered problem. In brief, the action is the translation of the previous phases in a concrete execution command in the target environment. In IoT environments, examples of actions include blocking of suspicious traffic, isolating of compromised system, and notification to final users or system administrators about security vulnerabilities.

The scenarios and the ways in which different stakeholders and functionalities interact with the ROBUST-6G platform will be explained in depth in the following sections. A preview of the stakeholders' roles and their contributions to the overall functionality and security of the platform is shown in Figure 2-3 which provides a visual depiction of the platform's components and their interactions with the IoT management features.

Figure 2-3 is composed mainly of five elementary blocks. The Zero-touch Security Management is the entry point of the system receiving requests from an external consumer and translating them in different steps. Once a new request is received the Security Management module requires some logic from the Trustworthy AI Services Layer for analysing the continuously monitored data provided by the Data Management Platform. This analysis, mainly AI/ML-driven, helps detecting the presence of anomalies.

Upon anomaly identification, the Decision step defines the resolution plan for mitigating them. Finally, the security management module enforces the action executing them in the target environment to erase the previously discovered anomalies. It is important to note, that in this idea of flexible and configurable closed loop, the function steps are preconfigured by different entities, but in the end, are executed by a security application which may stay even in the target infrastructure itself.

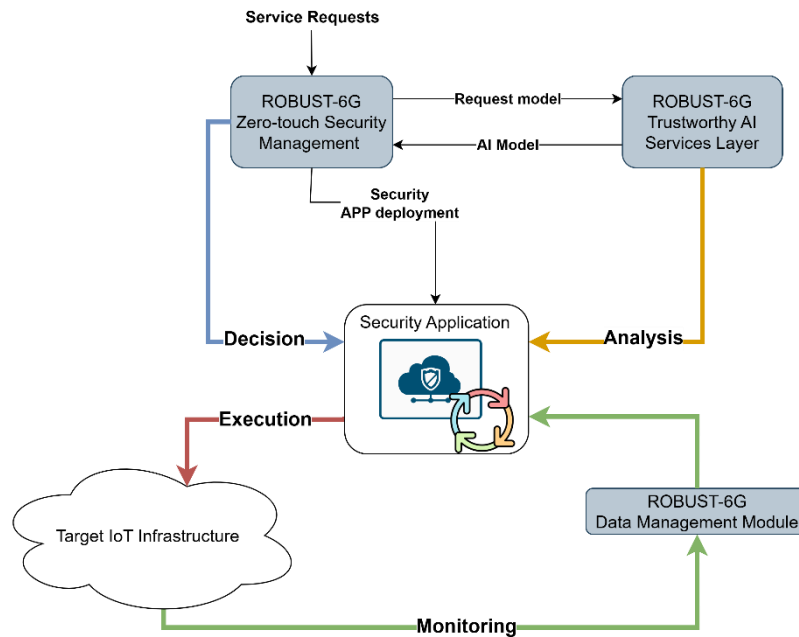


Figure 2-3: ROBUST-6G components interacting with the external world

Security automation provides significant value to stakeholders by reducing manual intervention. For Small and Medium Enterprises (SMEs), this approach offers competitive costs for implementing and maintaining robust security. Additionally, a flexible approach like this allows updating at runtime the closed loop functions, able to address last-minute changes.

2.2.2 Stakeholders definition, roles, and interactions

From the previous section, it is easy to identify key stakeholders interested in using such a platform. These stakeholders include users, SMEs, and Telco operators. This involvement lies mainly in the need for secure, efficient, and trustworthy IoT environments.

When referring to **common users**, it is possible to imagine people who manage their target environments, mainly composed of networks and IoT devices, such as smart homes or smart buildings using the ROBUST-6G platform. Nowadays, common users adjust settings, check device status, and receive alerts about possible security issues directly through the platform's interface. For instance, common users could be alerted to unauthorized access attempts or environment anomalies, enabling automatic or confirmative actions to solve the problem. In the scenarios proposed in the following sections, there is a deeper explanation of the interaction between common users and the platform.

Another example of stakeholders is SMEs. These businesses may use the ROBUST-6G platform to monitor device performance, improve operational efficiency, and reduce security threats in their IoT ecosystems. SMEs can operate in different sectors, including manufacturing, healthcare, and technology but they all have similar needs. For example, in a manufacturing setting, an attacker could exploit sensor vulnerabilities to manipulate production line outputs, resulting in product defects, material waste and generally in financial losses. To avoid such losses, it is important to apply security mechanisms that monitor and discover as early as possible such threats proposing a way to mitigate them. These kinds of examples are reported more in detail under the scenarios described in detail in Section 2.2.5.

In conclusion, iterating the previously defined concepts over bigger realities, telco operators may be considered as ROBUST-6G platform's stakeholders. Telecommunication operators provide infrastructure support, reliable network connectivity, data transmission and security standards to different platforms. Their participation is essential to preserve the trustworthiness and integrity of the communication routes. Furthermore, telco operators may interact with other stakeholders, such as businesses and end-users, to identify and address security issues in the access network or in the core network that may put device connectivity and network infrastructure at risk. For example, Telco Operators can detect unusual traffic reducible to DDoS attacks and can work with businesses to isolate the infected device or networks, to execute a contingency plan.

2.2.3 State-of-the-art for application in 6G networks

To design and develop a robust platform against threats and vulnerabilities in IoT contexts, cutting-edge technologies considering device characteristics must be incorporated. IoT devices operate with limited resources (energy, computational) which require lightweight security solutions. Hardware-based solutions, such as Key Agreement Protocols location-based or Physical RF fingerprinting functions, provide a secure path that can ensure device integrity and secure key storage. Similarly, lightweight cryptography has been designed to secure communications and data storage without exhausting device resources, making it suitable for IoT environments. This section aims to analyse and investigate previous researchers' solutions in the IoT field.

Networks and devices, present a wide range of well-known vulnerabilities, including weak authentication mechanisms, insecure communication channels, and obsolete encryption standards which expose IoT ecosystem to cyber-physical attacks. Taxonomies of these threats classify attacks into categories such as Denial of Service (DoS), Man-in-the-Middle (MITM), data tampering, and unauthorized access. For instance, weak encryption protocols may allow attackers to intercept and manipulate sensitive data, while insecure device configurations can provide entry points for unauthorized users. These vulnerabilities are related not only to IoT devices but also to automation applications and third-party applications such as IoT management platforms. Existing research studies provide hints for well-known threats and taxonomies related to smart home environments, and the entire IoT ecosystem, underlining the wide range of cyber-physical attacks [MHE+16, HLB+18, CZD23, BKT+22].

When dealing with 6G-enabled IoT environments other aspects may be taken into consideration for the correct design of a secure platform. The use of Deep Learning (DL) techniques for cyber-attack detection in IoT networks has been explored in recent studies like [JOR+23] and [GSS23]. DL algorithms may make a difference in the analysis step of a closed loop because they improve detection capabilities compared to traditional algorithms based on threshold or static feature analysis. For example, while traditional static algorithms rely on manual rule definition, DL models autonomously learn patterns from large datasets, improving the accuracy and making dynamic the anomaly detection. Furthermore, [KAK+23] developed a hybrid DL model combining Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) to detect IoT-specific attacks, achieving significant improvements in accuracy and false-positive rates. Similarly, [MKP+22] introduced a FL-based Intrusion Detection Systems (IDS) framework that protects data privacy while leveraging distributed DL models to enhance IoT security.

Another critical threat in IoT environments is cryptojacking, where attackers exploit the computational resources of IoT devices to mine cryptocurrencies. This attack type is particularly detrimental in resource-constrained IoT ecosystems, as it can drastically degrade device performance, increase energy consumption, and shorten device lifespans. Countermeasures against cryptojacking often involve anomaly detection methods, where DL-based systems can identify unusual patterns in resource usage and network behaviour indicative of mining activities. For example, [TAU22] proposed a lightweight DL-based model tailored for IoT devices to detect cryptojacking attacks with high accuracy and low computational overhead.

To summarise, there exist many studies on IoT security attacks in cyber-physical environments. Hardware-based solutions and lightweight cryptographic techniques provide robust security while maintaining efficiency. At the same time, DL solutions increase the ability to detect and respond to advanced threats, including cryptojacking and other sophisticated cyber-attacks. These studies may be considered as a valid background for the design and successive implementation of an efficient and secure ROBUST-6G platform with zero-touch security management against multiple cyber-physical threats in IoT environments.

2.2.4 Use case detailed description

This UC delves into the prediction, detection and mitigation of threats in IoT environments. It mainly analyses three scenarios which focus on providing proactive, reactive, and predictive security automation. Proactive security is the initial configuration necessary for setting up the target environment such that it is responsive and compliant with security mechanisms during its runtime. Reactive and predictive security instead considers all the aspects that follow during the runtime flow which may lead to the detection (in case it occurred) or prediction (in case it will occur) of threats. In addition, they also consider the relative mitigation steps. More details and workflows are described in Section 4.4.5.

Moreover, since the complexity of the problem is high, the use of well-known techniques may be efficient in searching for a solution. In these scenarios, in fact, it is planned to highlight the use of programmable closed-

loop functions. For the technical aspects related to the closed-loops, it may be beneficial to inspect deliverable D4.1, which elaborates more on this.

In the next section, including three subsections, we examine three different scenarios, with increased details. In all scenarios, the presence of IoT is in evidence, since the aim of the UC is to combine the meaningful information provided by the sensors spread across the environments and the network data. In particular, the first scenario describes a small company managed to reduce unnecessary energy costs caused by a compromised heating system thanks to advanced techniques that combine sensors and network data. The second scenario explores the device manipulation for a scope different than those for which they were designed from the production. For example, a compromised smart light can be used by an attacker for crypto mining purposes instead of illuminating a room. The third scenario, which is the last and most enhanced one, describes the relevant consequences of the compromised sensor causing a chain of reactions that might contaminate other external system of a different field. Fundamentally, the UC delves on the evil intent of altering devices to cause economic harm or gaining advantages from the fraudulent usage of IoT devices.

In the context of threat detection and mitigation in 6G-enabled IoT environments, it is important to define several KPIs. Among the different KPIs, the considered ones during the implementation of the proposed UC are:

- **Detection Accuracy:** With a target accuracy of 95%, it is expressed as the percentage of correctly identified threats (expressed as the ratio between true positives, false positives or true negatives and false negatives) among all the threats detected by the system.
- **Detection Time:** It is expressed as the amount of time that passes between the injection of an anomaly and its detection. In general, this is strongly related to the scenario complexity. Our target, compatible with the scenario, is of a maximum time of 2 min.
- **Mitigation Accuracy:** With a target accuracy of 95%, this is the percentage of proposed actions that are effectively carried out and leads the target environment to a correct behaviour mitigating the previously detected anomaly.
- **Mitigation Velocity:** Measured in terms of the number of closed loop function configurations needed to carry out mitigation actions. The valid target is the execution of the mitigation actions within a threshold lower or equals of three closed loops.
- **Mitigation Time:** The amount of time that intercourse between the problem detection and the full implementation of the corrective countermeasures. The goal time in this indicator is fewer than 10 minutes with additional consideration based on the scenario complexity.

The proposed KPIs are very important because with quantifiable values they indicate the efficiency of the proposed platform. In fact, it is very important to understand and measure how quickly, still accurately, the system can detect and mitigate threats. This is also a way to measure and compare different detection and mitigation strategies.

2.2.5 Scenarios

The following scenarios describe anomalies that UC2 aims to detect and mitigate, all set in a smart IoT environment, i.e., office, farm, home, where the communication is 5G/6G wireless. Such communication can be managed by a Telco operator, which offers ROBUST-6G security services for verticals (e.g., the company which owns the office in scenario 1. Nonetheless, the characteristics of the target environment i.e., standalone IoT, make it suitable to consider, per each scenario, a dedicated 5G/6G NPN (non-public network).

2.2.5.1 *Device violation to cause an economic harm (a)*

This scenario explores the world of small to medium-sized office environments, where there is potential security issues associated with the integration of IoT devices managed through a centralized IoT platform. An example is reported in Figure 2-4 depicting a standard office equipped with IoT devices (such as gateways, electricity meters, and heaters) managed together via a unique IoT platform. The possibility for bad actors to take advantage of the weaknesses of the devices and cause financial harm becomes real in this environment. For instance, it is easy to imagine a scenario in which a malevolent attacker obtains unapproved access to the IoT platform and makes an apparently benign command, such as “turn on” the office heater.

The platform request seems acceptable at first, but it raises concerns if it is being made on a non-working day. This is the core of the problem: the action accomplishes nothing useful for the company, instead it produces a waste of energy and causes financial loss that could have been avoided. However, the consequences may be

worse than energy waste such as equipment (servers, routers) damage due to the high temperature. Additionally, overheating could also damage the supply chains affecting the whole the whole company. Such cascading effects highlight the importance of a security platform and its efficiency in discovering and mitigating threats.

The described scenario is composed of a “single closed-loop” consisting of the well-known steps of collection, analysis, decision, and execution. At the bottom of Figure 2-4, the first step of collection is illustrated by device measurement monitoring including temperature readings, and network traffic. Through this continuous monitoring, and with the logic of an analysis step, any anomalies or unusual patterns are immediately identified guaranteeing a fast identification of possible threats. For example, unusual patterns are discovered mixing the high-temperature value in a room and the logs of the “turn on” command in the heating system. After the analysis, which could be threshold or AI-driven, the decision identifies the best remediation plan with necessary countermeasures to execute. These countermeasures could include putting the malicious actor on a blacklist, executing system shutdown commands, or releasing a software upgrade to fix vulnerabilities.

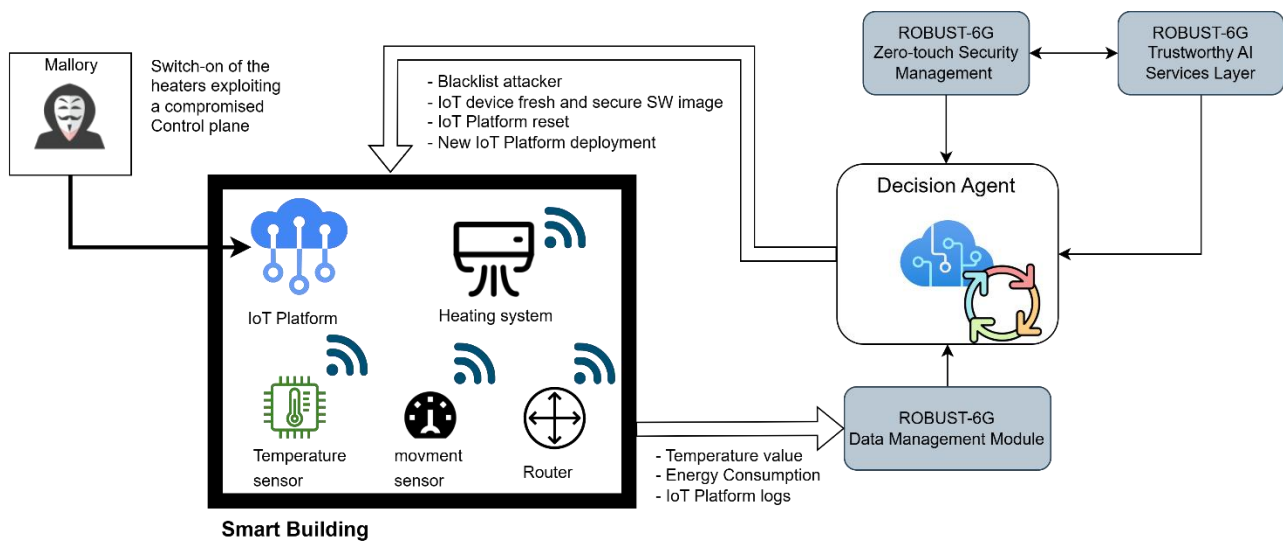


Figure 2-4: Device violation to cause an economic harm (a)

2.2.5.2 Fraudulent usage of device resources

This scenario leaves the floor to a smart building, where the presence of smart equipment creates a new channel for criminal activity. Imagine a hacker breaking into the system taking control of smart devices and then using their processing power for illegal activities like cryptocurrency mining (refer to Figure 2-5). This situation is tricky because the attack is hidden in the background. The compromised devices seem to work properly since they may be switched off and are not giving any alarm of being compromised. On the other side, a deeper analysis of the network-generated data helps in discovering the presence of the anomaly since the mining process may lead to intensive CPU consumption and network traffic.

The “two closed loops” at the centre of this storyline cooperate to identify and neutralise the cryptojacking threat. In the first loop, IoT device measurements are continuously monitored ensuring patterns deviations result in alarms or extra action to apply. The second loop starts by analysing network traffic to identify device behaviour anomalies. In this second case, the action execute is not explorative as in the first loop but is resolute (e.g., blacklisting attacker or device/IoT platform reset).

The people living in Smart Buildings depend on these devices for daily activities and their comfort. Additionally, network managers are relevant actors in this scenario keeping an eye on the infrastructure and protecting the integrity of the network itself. This enforces once more the need for a powerful system able to identify threats and define the most suitable remediation plan. Overall, the KPIs defined in Section 2.2.4, fit well in this scenario. Fast detection of compromised devices using unauthorised resources, and a quick reaction for threat identification and mitigation are fundamental criteria of validation for measuring the effectiveness of the proposed security platform.

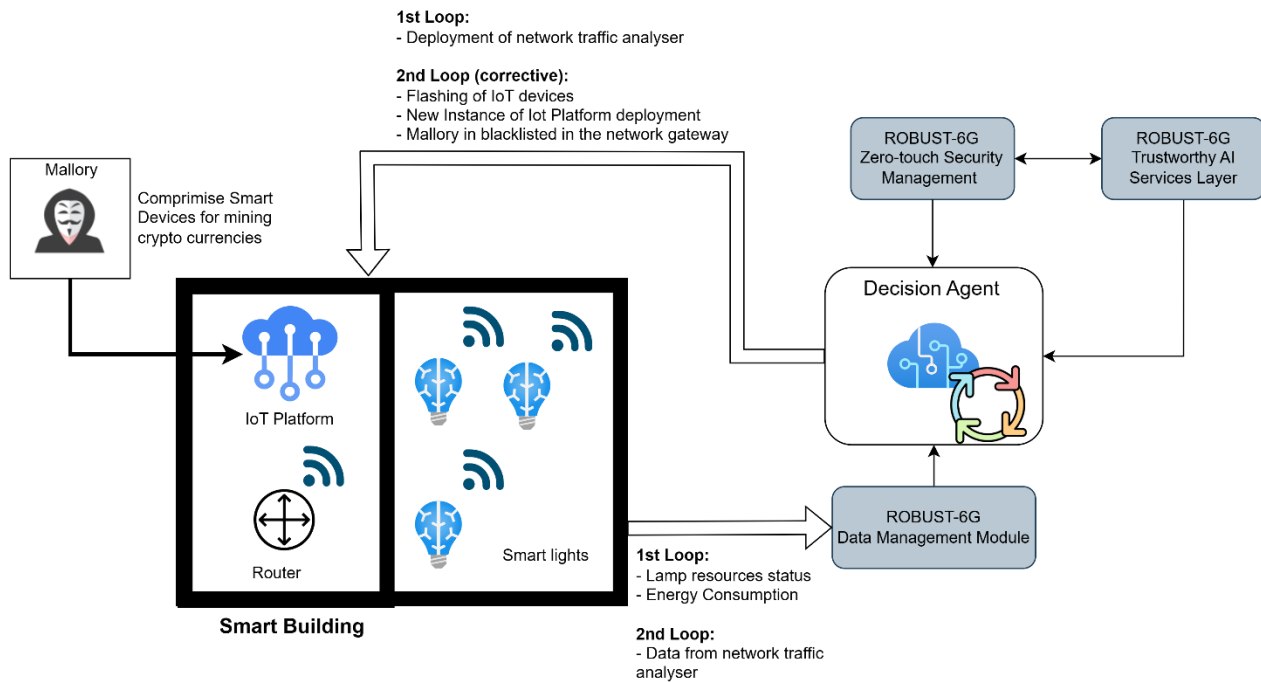


Figure 2-5: Fraudulent usage of device resources

2.2.5.3 Device violation to cause an economic harm (b)

The last scenario lies in the world of smart agriculture, where the attack surface exposure is quite vast. In this scenario, the financial damage from cyber-attacks could be more substantial and environmental aspects as waste of raw materials like water or fertilizers, need to be considered as well. This scenario investigates the influence of adjacent smart fields, exchanging data like temperature and humidity when one field becomes compromised. Vital resources like water, temperature and humidity are essentials for the correct lifecycle of plants. This evinces once more the importance of an efficient and secure mechanism to avoid damage from threat attacks.

The scenario depicted in Figure 2-6, reports the case where an attacker exploits sensor vulnerabilities in one field and manipulates the measurement data. This manipulation can cause cascading damage in the adjacent fields if they are based on automatic actuators, such as irrigators or heating systems, that depend on the information transmitted from the compromised sensors. To provide a more realistic example, suppose that a certain wheat requires a fixed percentage of humidity and a constant high temperature for growing in optimal conditions. By attacking a small (and weak) portion of sensors in the field and transmitting false values of temperature and humidity, an automatic irrigation/heating system based on such values can act incorrectly leading to wheat death and consequent significant financial losses.

The security solution proposed by ROBUST-6G implements several closed loops (intra and extra field) that deploy specific agents for cross-checking the correctness of the sensor values using as ground through external services like Open Weather Map. However, because sensors usually operate in low-power mode and transmit infrequently, alternative methods like RF fingerprint variations may be considered to detect attacks. In addition to data manipulation attacks, the system aims to address attacks concerning the physical layer like jamming attacks. In this case, appropriate solutions like frequency hopping, and beamforming using dMIMO can be applied to maintain reliable communications.

Overall, this scenario contains a hierarchy of several closed loops. The interconnection of these closed loops requires coordination between several sites and the usage of an external data sources like meteorological services to verify sensor readings.

To summarize, discovering anomalies requires real-time analysis of continuously monitored network traffic and sensor data as well as coordination across multiple local sites in case the data are combined as in FL. As mentioned in the previous scenario, among the expected KPIs it is important to underline also in this scenario the importance of the detection and mitigation time as well as the accuracy of the proposed mitigation action.

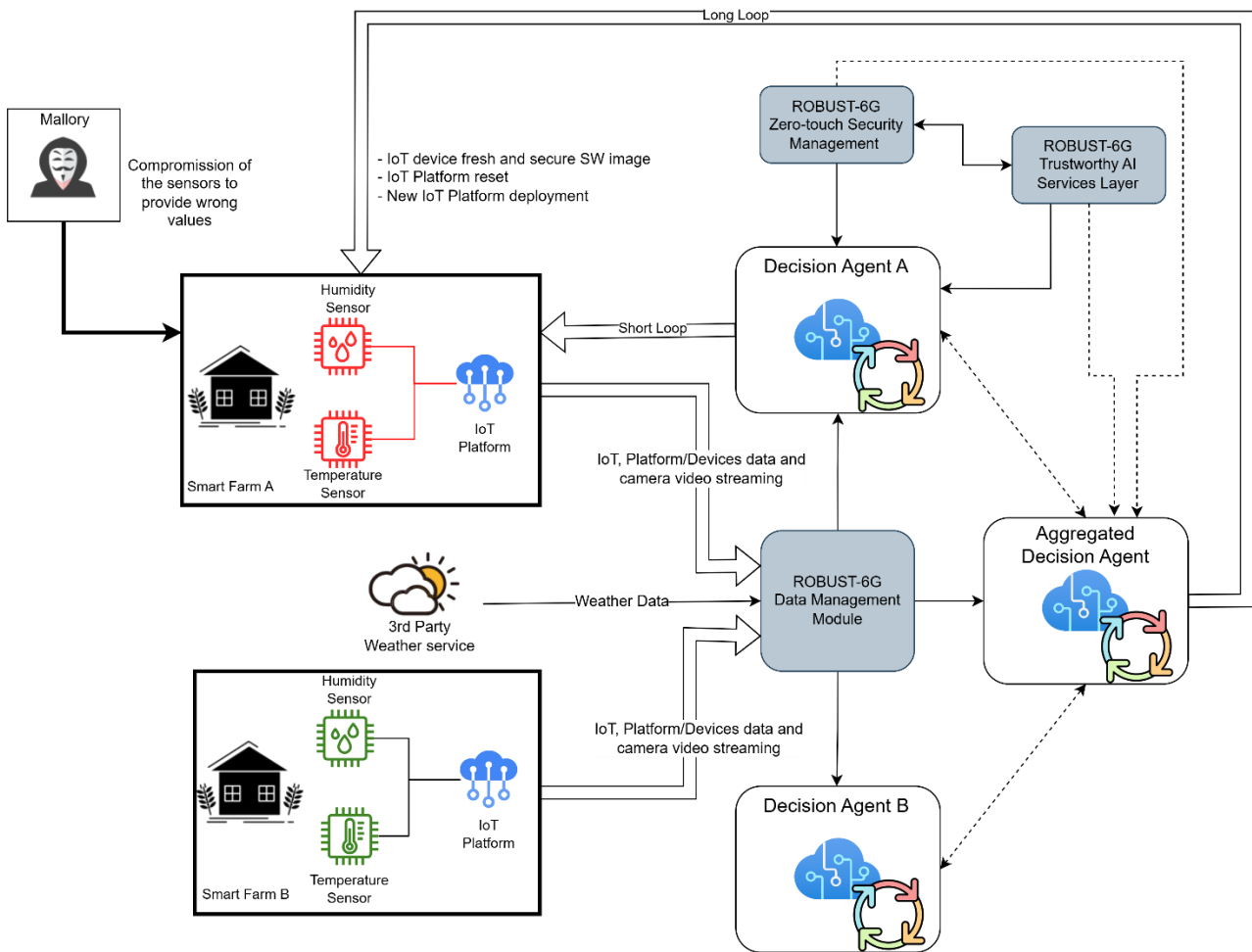


Figure 2-6: Device violation to cause an economic harm (b)

2.3 Security capabilities exposure with Network-Security-as-a-Service (NetSecaaS)

2.3.1 Motivation and overall description

Network-as-a-Service (NaaS) represents a cutting-edge approach to delivering network services, allowing third-party applications to seamlessly interact with the network through intuitive APIs. This innovative paradigm is poised to revolutionise the landscape of 6G technology, empowering application developers to harness network capabilities without requiring specialised network expertise. Spearheaded by the Global System for Mobile Association (GSMA) Open Gateway initiative, a framework has been devised to facilitate secure, on-demand, and controlled access to network functionalities via a standardised set of APIs. These APIs, developed under the CAMARA open-source project [Cam23], serve as an abstraction layer, shielding users from the intricacies of telecommunications while streamlining the utilisation of network features by external entities, such as application developers and enterprises.

This UC aims to expand the functionalities of the Open Gateway framework, enabling application developers and enterprises to seamlessly apply security policies by harnessing the enhanced capabilities of ROBUST-6G, termed as Network-Security-as-a-Service (NetSecaaS). The integration of ROBUST-6G with the Open Gateway framework, as illustrated in Figure 2-7, will be demonstrated.

As depicted in Figure 2-7, within the Open Gateway's northbound interface, novel APIs will be introduced to abstract security capabilities, building upon the ongoing development within the CAMARA project. These APIs will be meticulously crafted based on a thorough analysis of the security features offered by ROBUST-6G. On the southbound interface, the Data Fabric component, developed as part of the Data Management Platform, which is outlined in more detail below, will serve as the intermediary between ROBUST-6G and the Open Gateway. The Data Fabric will facilitate the exchange of security data from ROBUST-6G and the

ingestion of intent declaration data from the Open Gateway. These data flows will delineate the interaction between the core of the Open Gateway and the ROBUST-6G platform, taking advantage of the AI capabilities provided by ROBUST-6G to support an intent-based approach (the service API definition adopts a developer-friendly approach, abstracting away the inherent telco complexity found in network APIs). The *Semantic transformation & adaptation* block will be augmented to incorporate functionalities for mapping security intent declarations, modelled in CAMARA, into the data models employed by the ROBUST-6G Data Fabric. Simultaneously, the *Workflow engine* will be expanded to orchestrate interactions with ROBUST-6G, considering that a single data flow in the northbound interface may initiate multiple flows in the southbound interface.

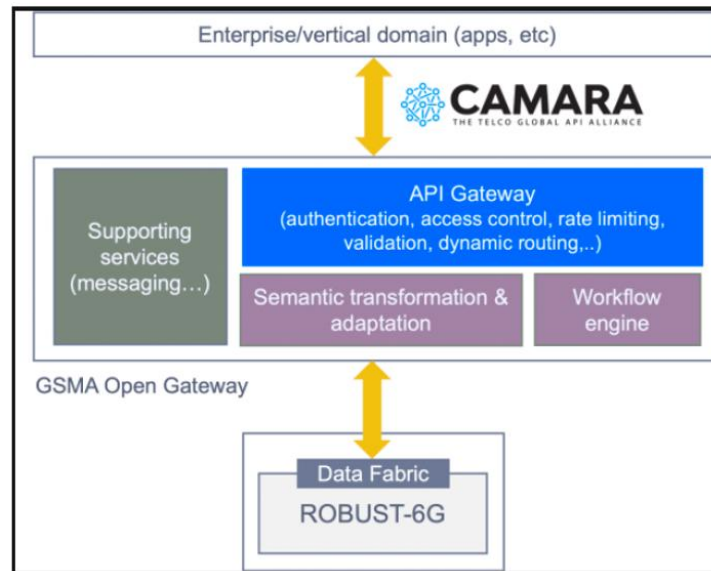


Figure 2-7: Integration of ROBUST-6G with Open Gateway

To validate the integration of ROBUST-6G with the Open Gateway, this UC will present tailored scenarios aimed at external users lacking expertise in network security. These users may include school network administrators or mobile application developers unfamiliar with security protocols. Users will express high-level security requirements, such as network encryption, layer 7-based filtering, policy scheduling, Internet Protocol Security (IPsec), parental control filtering, and key size considerations. Each scenario will be customised based on the resources of interest, ensuring that security measures align with user needs.

Prior to any interaction, authentication and authorisation will be conducted to establish a private and secure connection. These scenarios will demonstrate how ROBUST-6G capabilities can effectively address diverse security needs, showcasing the seamless integration of advanced security measures into the network infrastructure.

2.3.2 Stakeholders definition, roles, and interactions

In this UC we can meet the following main users and stakeholders:

- **Application Developers:** They are the primary users of the NaaS platform. Their role involves utilising the NaaS APIs provided by the Open Gateway framework to integrate network functionalities into their applications without needing extensive networks expertise. In this UC, they will also interact with the enhanced security capabilities offered by ROBUST-6G through the NetSecaaS component.
- **Enterprises:** They are another key stakeholder group that benefits from NaaS and NetSecaaS. They can leverage these services to enhance the security posture of their networks and applications without the need for specialised security knowledge.
- **Network Administrators:** They play a role in configuring and managing the Open Gateway framework within their organisation's network infrastructure. They ensure that the integration with ROBUST-6G is seamless and that the security policies defined by application developers and enterprises are effectively implemented. For instance, network administrators can enable network encryption to ensure that all data traffic within the network is protected from unauthorized access, meeting the organization's security and privacy requirements.

- **End Users:** Though they do not directly interact with the NaaS or NetSecaaS platforms, end users such as parents and general consumers benefit indirectly from the security features enabled by ROBUST-6G. For instance, parents may utilize applications that offer features like layer 7-based filtering for child safety, allowing parents to filter content or set usage policies without needing technical skills.
- **Regulatory Bodies:** These entities ensure that ROBUST-6G's security functionalities meet established standards and regulations. They play a key role in overseeing and guiding compliance with data protection, privacy, and encryption requirements.

By aligning the roles and interactions of users/stakeholders with various functional components of the platform, the integration of ROBUST-6G with the Open Gateway framework for NetSecaaS can be effectively demonstrated and validated, showcasing its practical applicability and benefits to users with diverse skill sets.

2.3.3 State-of-the-art for application in 6G networks

2.3.3.1 Open Gateway

NaaS represents a major shift where network operators can offer telecommunication capabilities for external use through APIs, enabling easier access to monitoring and configuration functions. These APIs empower telco networks to become programmable service platforms accessible to developers, Application Service Providers (ASPs), and enterprises, fostering seamless integration of applications with the network.

As highlighted in the 6G-DATADRIVEN project [Dat24], the development of NaaS requires a collaborative effort, bringing together telco standard bodies, IT/cloud communities, industry associations, and open-source projects. This collaborative environment necessitates a clear delineation of responsibilities to prevent redundancy and fragmentation within the NaaS ecosystem.

To address this need, GSMA launched the Open Gateway initiative at the Mobile World Congress (MWC) Barcelona in 2023. The mission of GSMA Open Gateway is twofold: i) to establish a governance framework for NaaS, covering both technical and business aspects; and ii) to secure operator commitment to launching universal NaaS API services by 2023.

The Open Gateway initiative acknowledges the foundational work done by three key organisations:

- **Linux Foundation's CAMARA:** This organisation focuses on the exposure aspect, defining user-friendly and open APIs for external consumption. CAMARA hosts and manages these APIs, ensuring they meet service and business needs while adhering to the Apache2.0 license.
- **GSMA:** GSMA plays a pivotal role in defining the technical and business aspects of NaaS. It specifies how third party-facing APIs are supported by telco capabilities and establishes agreement templates for federation between operator networks and third parties. GSMA oversees technical and business workstreams through the Operator Platform Group (OPG), Operator Platform API Group (OPAG), Open Gateway Technical Stream (OGWTS), and Wholesale Agreement Services (WAS) groups.
- **TM Forum:** TM Forum addresses the operational aspect, ensuring efficient management and operation of third party-facing APIs. It focuses on operational functionality provided by Operational Support Systems (OSS), Business Support Systems (BSS), and online charging systems under the Open Digital Architecture (ODA).

Figure 2-8 depicts the contributions made by these organizations within the Open Gateway actor-role model. The **Consumer**, comprising developers, ASPs, Independent Software Vendors (ISVs), and Enterprise customers, generates code that interacts with the APIs. On the other hand, the **Aggregator**, which could be a hyperscaler/OTT (Over-The-Top) or an operator, acts as a sales representative for the Open Gateway community. Its effectiveness increases when it represents numerous operators.

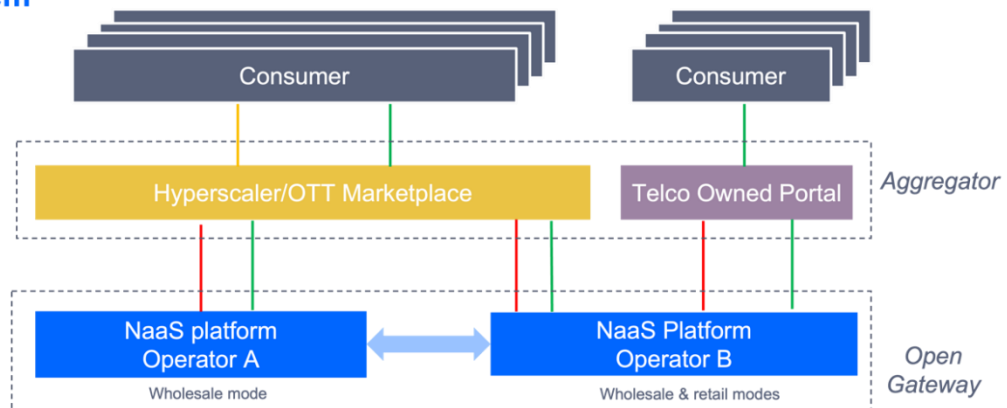
While each **Operator** establishes its own Terms and Conditions with the channels, alignment on product (standard APIs) and business framework (agreement templates, charging models) is essential. This synergy ensures a cohesive ecosystem where operators and channels operate seamlessly within predefined standards and agreements.

NaaS ecosystem

The **Consumer** is a Developer, Application Service Provider (ASP), ISV, Enterprise customer, that creates code that invokes the APIs

The **Aggregator** may be an hyperscaler/OTT or an operator. It sells on behalf of the Open Gateway community and is effective when it represents a high number of operators.

Each **Operator** sets its own T&Cs with the channels, but there needs to be full alignment on product (standard APIs) and business framework (agreement templates, charging models)










Industry forum	Legend	Scope	Description
		Interconnection & Agreements	Solutions from technical, product and business standpoint to ensure cross-operator consistency
		Service API	QoD, Device Location, Device Status, SIM Swap, OTP Validation, Carrier Billing..
		Service Mgmt API	Service execution validation, service status, service consumption, service ticketing
		Operate APIs	Service LCM, Developer/customer/merchant LCM, assurance, billing.
		Enhanced CAMARA APIs	Hyperscaler/OTT may use CAMARA APIs (service & service mgmt APIs) to create own enriched products

Figure 2-8: NaaS ecosystem, roles and usage of APIs

On one hand, GSMA's focus is squarely within the telecommunications domain, where it sets out the essential capabilities that all operators must provide for third-party access to ensure global reach and scalability. These essential capabilities, termed Open Gateway services, are prescribed by GSMA. Additionally, GSMA is tasked with prioritising and managing the roadmap of Open Gateway services, aligning them with market demands and the readiness of underlying technologies. Moreover, GSMA architects the platform that individual operators utilise to implement and expose these Open Gateway services.

On the other hand, CAMARA and TM Forum concentrate on the APIs enabling programmatic access to Open Gateway services. These APIs are categorised into three groups:

- **Service APIs:** These APIs facilitate the invocation of specific Open Gateway services tailored for various applications. Examples include Quality on Demand (QoD) API, Device Location API, and One-Time Password (OTP) validation API, each providing application-specific functionality.
- **Service Management APIs:** These APIs enable management actions within Open Gateway services, such as ordering activation/deactivation of functionalities, monitoring, eligibility checks, and consumption verification.
- **Operate APIs:** These APIs offer transversal functionality necessary to commercialise Open Gateway services, ensuring their operability and monetisation. Functions provided by Operate APIs include registration and onboarding of third parties, service fulfilment (e.g., provisioning, activation, modification), service assurance (e.g., incident management, performance monitoring), and billing.

In terms of API ownership, CAMARA is responsible for defining, developing, testing, and maintaining Service and Service Management APIs, while TM Forum oversees Operate APIs.

Regarding targeted consumers, CAMARA APIs are utilised by third parties in their applications either through aggregators (wholesale model) or via telco portals (retail model). Aggregators may develop and expose the additional Enriched APIs by adapting or combining CAMARA APIs. However, Operate APIs are not accessible to third parties; they are primarily used for integration with aggregators and portals.

For further information on CAMARA APIs and Operate APIs, refer to the White Paper published in [Gsm23].

2.3.3.2 Architectural approaches

The notion of capability exposure has garnered significant attention across various projects, especially following the introduction of public-private networks in the 3rd Generation Partnership Project (3GPP) Rel-16 specification. This concept, also known as Public Network Integrated Non-Public Network (PNI-NPN), facilitates the provision and operation of End-to-End (E2E) services across multiple administrative domains, merging resources from both Public Land Mobile Networks (PLMNs) and private (on-premises) networks.

As already introduced in the 6G-CHRONOS project [Chr24], projects such as 5G-VINNI (ICT-17), 5Growth (ICT-19), 5G-CLARITY (ICT-20), and Hexa-X (ICT-52) have extensively explored capability exposure, each adopting distinct architectural approaches tailored to their specific scopes and targeted UCs. Notably, these projects have provided valuable insights into this domain, which serve as foundations for subsequent endeavours.

For instance, the 5G-VINNI project focused on establishing an E2E 5G experimentation facility accessible to vertical industries, emphasising network slicing. This experimentation introduced a level-based framework for capability exposure, delineating four levels wherein tenants could access different operational capabilities from slice providers. Meanwhile, 5Growth aimed to validate 5G-enabled vertical applications using facilities like 5G-VINNI, primarily opting for exposure level 1 to access the experimentation facility while leveraging their own on-premises infrastructure for sensitive systems.

In a similar vein, the 5G-CLARITY project concentrated on designing a system offering diverse capabilities within private industrial network environments, defining the Mediation Function to regulate access control for consumers, including Mobile Network Operators (MNOs) and hyperscalers. This function served as a single-entry point, ensuring granular access control over provider-managed resources.

Furthermore, the Hexa-X project expanded on these concepts by introducing API Management Exposure, refining the solutions proposed by 5G-CLARITY. Notably, it addressed the dynamic nature of resources, particularly at the extreme edge of the computing continuum, and included software/application developers as tenants to integrate their Continuous Integration and Continuous Delivery/Deployment (CI/CD) pipelines into operators' systems.

Building upon the expertise garnered from these projects, the 6G-CHRONOS project as another significant endeavour, focusing exclusively on security-centric capability exposure. By leveraging an API gateway, it enhances integration with Operational Technology (OT) systems while offering improved control over access, QoS, network rate distribution, and protection against DDoS attacks. Additionally, the API gateway facilitates the monitoring of key metrics and logs, enabling early detection of communication flow anomalies.

The integration of ROBUST-6G with the Open Gateway framework represents a significant advancement in the exposure of network security capabilities, particularly in the context of NetSecaaS. This innovative approach introduces advanced APIs into the Open Gateway's northbound interface, effectively abstracting the sophisticated security functionalities of ROBUST-6G. These new APIs provide a streamlined, intuitive, and user-friendly interface, empowering application developers and enterprises to seamlessly access and deploy robust security features without requiring in-depth knowledge of complex network security protocols. This ease of access to advanced security differentiates ROBUST-6G from previous projects, offering a cutting-edge solution to meet modern security needs.

2.3.4 Use case detailed description

The UC 3 system architecture is pictured in Figure 2-9. This system architecture consists of the main following components:

Integration Layer: It is a crucial component in the transition towards NetSecaaS, deploys a range of critical capabilities to mediate interactions between applications and telecommunications resources, ensuring seamless and secure integration between the network and applications. This layer comprises two main components:

1. **Exposure Gateway:** This gateway provides all necessary capabilities to manage the interaction between the stakeholders, with a strong focus on security. This includes the publication and discovery of service APIs, access control (authentication and authorisation of applications), auditing, accounting, and logging. Through these functionalities, a secure and controlled environment for interactions between the network and applications is ensured.
2. **Transformation Function:** This function plays a crucial role in maintaining and executing security mappings between developer-friendly service APIs and low-level security APIs. It aims to ensure that interactions between stakeholders and network security resources adhere to established security standards. This is achieved through the execution of workflows designed to enforce security policies, ensuring that communications are secure and reliable.

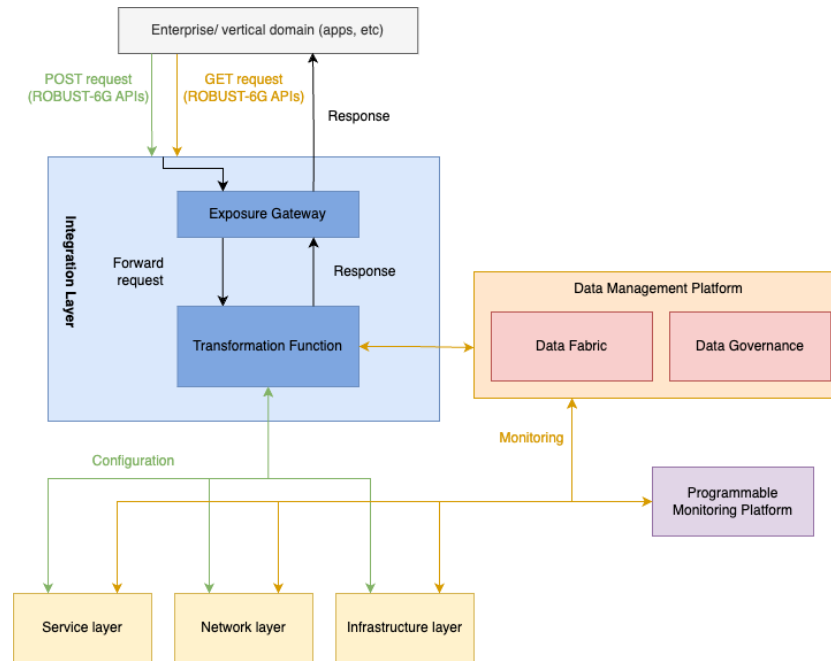


Figure 2-9: Use Case 3 system architecture

Data Fabric: This platform serves as a self-contained data ecosystem, facilitating seamless data integration and manipulation according to the principles outlined in the data mesh paradigm [Deh22]. Its primary function is to organise data into source-aligned domains, equipped with intuitive mechanisms for efficiently constructing and disseminating data products across various consuming data domains. This platform is pivotal for data manipulation and is slated for the deployment explained in Section 5.

Data Governance: Federated data governance mechanisms for guaranteeing high-quality data, privacy, and secure access to data as defined by owners of data domains.

2.3.5 Main scenario description

Use Case 3 focuses on demonstrating the integration of ROBUST-6G security capabilities into third-party applications via the Open Gateway framework. This integration will be showcased through a Proof of Concept (PoC) deployed in the 5TONIC lab [Ton24].

For this UC, the set of KPIs with which to evaluate the performance and effectiveness of the proposed system is determined below.

- API call average latency of 300ms and max latency of 1s for external applications waiting for an answer from the Open Gateway API.
- API CPU usage below 30% as part of the API responsiveness.
- At least 50% of security capabilities implemented by ROBUST-6G are exposed through standard CAMARA APIs.

To achieve these KPIs, the focus will be on streamlining processes and infrastructure to minimize latency and CPU usage. Collaboration with ROBUST-6G developers will ensure seamless integration of at least 50% of security capabilities into standard CAMARA APIs. Additionally, efforts will focus on optimizing resource usage, with a target of reducing overall consumption by 30%.

The UC scenarios tailored to users lacking expertise in network security, such as school network administrators or mobile application developers, demonstrate the practical applicability of NetSecaaS. By addressing high-level security requirements expressed by these users, the integration validates the seamless integration and effectiveness of advanced security measures provided by ROBUST-6G.

3 ROBUST-6G requirements

This section contains the set of requirements that are defined for ROBUST-6G. The selection and description of the different requirements have been structured in different groups according to the functional and non-functional capabilities expected in the different elements, which will be part of the ROBUST-6G architecture that is described in detail in Section 5.

Specifically, the distribution of requirements has been done based on the four main application domains of the project, namely: Physical Layer Security, Data Management, AI-Driven Distributed Security and Zero Touch Security Management. In addition, as a first initial building block, a few Global Requirements are listed that are crosscutting to the four application domains mentioned above.

In this section to define the system requirements, we adopt a compact tabular format to describe and report the requirements associated with any of the four application domains and the global requirements block. More specifically, Table 3-1 shows the structure of the requirements table.

Table 3-1: Structure of the tables with requirements

Req. ID	Title	Requirement description	Type	Origin
Application Domain [<i>X: Domain description</i>]				
[<i>RX.Y</i>]	[<i>Group title</i>]	[<i>Description</i>]	[<i>Type</i>]	[<i>Origin</i>]

Where the different attributes of the requirements are as follows:

- [*X: Domain description*]: To facilitate the placement of each requirement, they are grouped into different application domains. Each domain is labelled by an X digit.
- [*RX.Y*]: A unique identifier for each requirement, which can be used to reference them in a simpler way. It has the structure RX.Y, where X is the application domain number and Y is a sequential number of each requirement within the application domain, starting with 1.
- [*Group title*]: A very short indicator of the main characteristic to which the requirement is associated, thus being able to quickly identify them among the different requirements of the application domain.
- [*Description*]: It contains a brief, but complete, description of the requirement. For each, the level of the requirement is underlined with respect to its priority, indicating if it is mandatory, recommended or optional. To be more specific, we follow the keywords defined in RFC 2119 [Bra97] to indicate requirement levels:
 - MUST and SHALL mean a mandatory requirement that needs to be fulfilled.
 - SHOULD denotes a recommended requirement, which can be ignored with valid reasons in particular circumstances, but the full implications must be understood and carefully weighed before choosing a different course.
- [*Type*]: Clear identification of the type of requirement, which can then be classified according to one or more of the types described below (together with its abbreviation):
 - Functional (**Func**): It is focused on service or system behaviours, activities, or task. It may also cover actions related to component management.
 - Non-Functional (**NFunc**): It is a qualitative feature such as performance, security, reliability, portability, usability, etc.
 - Technical (**Tech**): It contains details related to the technology stack, integration, infrastructure stuff, plans for scaling up (*vertical*) or down (*horizontal*), data formats, APIs, etc.
 - Operational (**Oper**): It refers to the deployment, monitoring, logging, maintenance activities, among others. It may address deployment strategies, environment specifications, continuous integration and delivery pipelines.
 - Business (**Biz**): Business characteristics of a proposed system from the viewpoint of the system's stakeholders.
 - User (**User**): It is for potential consumers of ROBUST-6G.
- [*Origin*]: Brief indication of the origin of the requirement. This may be the baseline of the project scope as stated in the DoA, a specific UC or a particular element within the ROBUST-6G system.

Starting with the requirements providing key functionalities for the ROBUST-6G system, a pool of global requirements is shown in Table 3-2. They are breaking them down later into specific functionalities with respect to the infrastructure supporting the ROBUST-6G system and all its operations, and finally detailing the qualitative features to be considered. That is, these functionalities are elaborated upon by outlining the qualitative attributes and performance characteristics that must be taken into consideration to ensure seamless integration and optimal performance of the system.

Table 3-2: Overall system requirements for the ROBUST-6G system

Req. ID	Title	Requirement description	Type	Origin
Application Domain 0: GLOBAL REQUIREMENTS				
R0.1	Security capabilities	The ROBUST-6G system <u>must</u> provide secure, privacy-preserving, reliable, resilient, accountable, trustworthy, and sustainable capabilities.	Oper, NFunc, Biz	DoA
R0.2	Component interactions	The ROBUST-6G Security Orchestrator <u>should</u> interact with the Data Management Platform and with the Physical Layer.	Oper, Func	DoA
R0.3	Security capabilities	The ROBUST-6G system <u>must</u> implement observation, analysis, detection of threats and reaction to threats, as well as alert generation.	Oper, Func	DoA
R0.4	Type of closed-loop	The ROBUST-6G system <u>should</u> support a programmable approach using dynamic closed-loops.	Oper, Func	DoA
R0.5	6G layers	The ROBUST-6G system <u>should</u> be able to work on the 6G system layers defined as service, network and infrastructure.	Oper, Func, Biz	DoA
R0.6	Devices and environments	The ROBUST-6G system <u>must</u> be able to monitor metrics of devices located in the Edge and Far-Edge environments, especially considering IoT devices.	Tech, Func	DoA, UC2
R0.7	Threat detection	The ROBUST-6G system <u>should</u> be able to detect incidents based on rule-based or AI-driven mechanisms.	Oper, Func	UC2, DoA
R0.8	Type of closed-loop	The ROBUST-6G system <u>must</u> handle reactive/predictive closed-loops automatically.	Tech, Func	DoA
R0.9	Environments	The ROBUST-6G system <u>should</u> strive to cover the cloud-edge continuum in a decentralized environment, accommodating the capabilities of current technologies.	Oper, NFunc, Biz	DoA
R0.10	Virtualization	The ROBUST-6G system <u>should</u> be virtualised.	Tech, NFunc	DoA
R0.11	Security guarantees	The ROBUST-6G system <u>should</u> ensure robust mechanisms of timing accuracy, fairness, and privacy in detection and mitigation processes.	Oper, NFunc, Biz	DoA, All UCs
R0.12	Energy reduction	The ROBUST-6G system <u>should</u> use an energy-aware approach to reduce the energy consumption.	Tech, NFunc, Biz	DoA
R0.13	Privacy preservation	The ROBUST-6G system <u>shall</u> deal with the privacy threats that could compromise its efficiency and functioning. Privacy-enhancing mechanisms should be an integral part of the overall platform.	Tech, NFunc, Biz	All UCs

Table 3-3 shows the ROBUST-6G system requirements covering the main functional solutions to address security, privacy, authentication and anomaly detection at the physical layer of the infrastructure using advanced techniques such as ML, FL and PLS.

This following list also adds some non-functional features focused on robustness, energy efficiency, privacy by design and attack mitigation, with the aim of optimising security in 6G networks.

Table 3-3: Requirements of the physical layer security in the ROBUST-6G system

Req. ID	Title	Requirement description	Type	Origin
Application Domain 1: PHYSICAL LAYER SECURITY				
R1.1	PHY security technologies	The ROBUST-6G system <u>must</u> provide PLS based security schemes for 6G leveraging massive Multi-Input Multi-Output (mMIMO), RIS, dMIMO.	Oper, Func	DoA
R1.2	Authentication and key agreement	The ROBUST-6G system <u>must</u> provide low latency and low footprint authentication and key agreement protocols for the considered UCs.	Oper, Func	DoA, UC1
R1.3	Fake base station identification	The ROBUST-6G system <u>must</u> include a technique for the identification of false base stations.	Oper, Func	DoA
R1.4	Localization privacy at the PHY	The ROBUST-6G system <u>must</u> provide physical-layer based solutions for localization privacy.	Oper, Func	DoA
R1.5	Trustworthy sensing at the PHY	The ROBUST-6G system <u>must</u> provide physical-layer based solutions for trustworthy sensing.	Oper, Func	DoA
R1.6	Anomaly detection at the PHY	The ROBUST-6G system <u>must</u> provide physical-layer based solutions for generalized anomaly detection.	Oper, Func	DoA
R1.7	Online attacker mitigation	The ROBUST-6G system <u>should</u> include online attack and attacker identification and mitigation solutions with the help of online AI/ML learning mechanisms.	Oper, Func	DoA
R1.8	Self-devices configuration	The ROBUST-6G system <u>should</u> include ML solutions to learn how to configure the devices and what signals they should transmit to improve the confidentiality of transmissions using wiretap coding.	Oper, Func	DoA
R1.9	Joint resource optimization and confidentiality	The ROBUST-6G system <u>should</u> integrate confidentiality solutions with authentication and Secret Key Generation (SKG) to optimize the resources (in terms of energy consumption, but also communication overhead).	Oper, Func	DoA
R1.10	Challenge response-based authentication	The ROBUST-6G system <u>should</u> include solutions for authentication based on challenge-response approach operating at the PHY.	Oper, Func	DoA
R1.11	Robust SKG techniques	The ROBUST-6G system <u>should</u> include solutions for the SKG techniques robust against eavesdropping, injection (man-in-the-middle), spoofing, and jamming.	Oper, Func	DoA
R1.12	ML models protection	The ROBUST-6G system <u>should</u> be robust against adversarial attacks against ML models used for PHY security.	Oper, NFunc	DoA
R1.13	Privacy-preserving	The ROBUST-6G system <u>should</u> include privacy-preserving solutions while following principles such as privacy by design, local processing, and confidential computing, as well as anonymization, pseudonymization, obfuscation, and perturbation.	Oper, NFunc	DoA
R1.14	Localization privacy-preserving	The ROBUST-6G system <u>should</u> include solutions for positioning privacy at the PHY using channel charting.	Oper, Func	DoA
R1.15	Image forensics-based anomaly detection	The ROBUST-6G system <u>should</u> include anomaly detection and restoration techniques inspired to image forensics based on an image of the environment obtained from both in-band and opportunity signals.	Oper, Func	DoA

R1.16	General cross-layer anomaly detection	The ROBUST-6G system <u>should</u> include generalized cross-layer anomaly detection techniques using continuous learning and unsupervised learning.	Oper, Func	DoA
R1.17	Federated schemes	The ROBUST-6G system <u>should</u> include federated solutions operating across several devices (both users and network components) for spatial correlations in detecting federated attacks (e.g., jamming covering an area or distributed attacks).	Oper, Func	DoA
R1.18	Creation of PHY attack database	The ROBUST-6G system <u>should</u> create a database of attacks at the physical layer (PHY) and sensing, generated by contributions from all partners participating in the PLS-related WP.	Tech	DoA
R1.19	AI-enabled RF fingerprint library	The ROBUST-6G system <u>should</u> develop an AI-enabled library of known RF fingerprints for different identified attacks.	Oper, NFunc	DoA
R1.20	Threat localization using AoA and CSI	The ROBUST-6G system <u>should</u> leverage the Angle of Arrival (AoA) and Channel State Information (CSI) to help ML models estimate the location of threats with less training data.	Tech	DoA
R1.21	RF fingerprint migration for seamless IoT communication	The ROBUST-6G system <u>should</u> support the adaptive migration of RF fingerprints among base stations in smart city environments to enable seamless and secure communication for IoT devices across different network nodes.	Tech, Func	DoA
R1.22	Predictive models for RF fingerprint adaptation	The ROBUST-6G system <u>should</u> develop predictive models to anticipate changes in RF fingerprints for low-power, infrequently communicating IoT sensors, enabling privacy-preserving and robust sensing.	Tech, NFunc	DoA

Table 3-4 illustrates the requirements associated with data management in the ROBUST-6G system. This list of requirements is shown based on a first exposure related to APIs and security in external access to managed information, moving on to review data security and governance, including the necessary policies related to the management of sensitive data. This is followed by the collection and discovery of data from different sources for further data management, with the aim of establishing a robust infrastructure. Finally, the monitoring requirements are outlined to further maintain the advanced security capabilities of the system.

Table 3-4: Data management requirements in the ROBUST-6G system

Req. ID	Title	Requirement description	Type	Origin
Application Domain 2: DATA MANAGEMENT				
R2.1	External access APIs	The ROBUST-6G system <u>shall</u> define user-friendly APIs for external consumers to gain access to exposed capabilities.	Tech, Func, Biz, User	UC3
R2.2	API Security	The ROBUST-6G system <u>shall</u> provide secure API access for external consumers to the internals of transformation mapping between the service API and the network API.	Tech, Func, Biz, User	UC3
R2.3	External APIs discovery	The ROBUST-6G system <u>shall</u> have mechanisms to make these APIs discoverable to external consumers.	Tech, Func, User	UC3
R2.4	External APIs	The ROBUST-6G system <u>should</u> provide information to the outside through APIs designed for secure information transaction.	Oper, Func	DoA
R2.5	Data access authorization	The ROBUST-6G system <u>shall</u> allow access to data only to authorised data consumers based on the permissions defined by data owners.	Tech, Func, User	Dataspace

R2.6	Data access authentication	The ROBUST-6G system <u>shall</u> include authentication mechanisms for accessing data.	Tech, Func	Dataspace
R2.7	Data privacy labelling	Defining data governance policies for data access <u>shall</u> account for sensitive data.	Oper, NFunc, Biz	Data Governance
R2.8	Data provenance	Tracing the data's history throughout its life cycle <u>should</u> be necessary for instilling trust in the data.	Tech, NFunc, Biz	Data Fabric
R2.9	Batch data sources	The ROBUST-6G system <u>must</u> support mechanisms for collecting data in batch mode.	Tech, Func	Data Fabric
R2.10	Streaming data sources	The ROBUST-6G system <u>must</u> support mechanisms for collecting data from streaming data sources.	Tech, Func	Data Fabric
R2.11	Data cataloguing	Data consumers within ROBUST-6G <u>should</u> require a means to discover available data.	Tech, Func, User	Data Governance
R2.12	Distributed data management	Data management <u>should</u> be distributed to ensure scalability and adaptability in dynamic data exchange scenarios.	Tech, NFunc, Biz	Data Fabric
R2.13	Data product ownership	Data owners <u>must</u> be accountable for the data products created and exposed with the Data Fabric.	Tech, NFunc	Data Fabric
R2.14	Heterogeneous data	The ROBUST-6G system <u>must</u> enable the integration of heterogeneous data from data sources of different types.	Oper, NFunc, Biz	Data Fabric, PMP
R2.15	Duplicity of the information	The ROBUST-6G system <u>shall</u> avoid duplicity of information, making efficient use of resources.	Func	
R2.16	Monitoring	The ROBUST-6G system <u>shall</u> support the monitoring of network and security resources.	Oper, Func	UC3
R2.17	Data monitoring	The ROBUST-6G systems <u>should</u> collect data across different layers of the 6G system: service, network, and infrastructure.	Oper, Func	DoA, PMP
R2.18	Data monitoring	The ROBUST-6G system <u>should</u> integrate tailored monitoring agents for far-edge and edge monitoring for selective distribution of monitoring data.	Oper, Func	DoA, PMP
R2.19	Aggregation and correlation	The ROBUST-6G system <u>should</u> have a correlation mechanism for similar data collected from different environments.	Tech, Func	DoA
R2.20	Extending capabilities	The ROBUST-6G system <u>should</u> support new monitoring modules/tools to extend the capabilities without modifying the core of the platform.	Tech, Func	
R2.21	New health metrics	The ROBUST-6G system <u>should</u> be able to create new metrics from those previously monitored.	Tech, Func	DoA
R2.22	Threat detection	The ROBUST-6G system <u>should</u> aggregate information to preprocess early threat detection.	Oper, Func	DoA
R2.23	Agents' reconfiguration	The ROBUST-6G system <u>should</u> be able to reconfigure the monitoring agents dynamically.	Tech, Func	DoA
R2.24	Communication security	The ROBUST-6G system <u>should</u> have secure communication between its modules, with special emphasis on the transmission of the agents with the module in charge of aggregating and preprocessing the information.	Tech, NFunc, Biz	

Next, Table 3-5 presents a series of security requirements based on distributed AI that must be considered in the ROBUST-6G system for its correct operation. Firstly, those related to the strategic objectives of this application domain are introduced; and aligned with the goals and priorities of the stakeholders. Subsequently, it is shown the requirements related to the capabilities needed to increase security and privacy to ensure that the ROBUST-6G system is secure and protects sensitive data are listed. This is followed by requirements on

the specific technical capabilities of the system, which implement key functionalities, while finally underlining the system's commitment to sustainable and responsible practices.

Table 3-5: Distributed AI-driven security requirements in the ROBUST-6G system

Req. ID	Title	Requirement description	Type	Origin
Application Domain 3: DISTRIBUTED AI-DRIVEN SECURITY				
R3.1	AI/ML ethics guidelines	The ROBUST-6G system <u>should</u> align with international AI/ML ethics guidelines to ensure ethical considerations are embedded in the development and deployment processes.	Oper, NFunc, Biz	DoA, UC1
R3.2	Green scheduling	The ROBUST-6G decentralized learning framework <u>should</u> integrate solutions to reduce carbon emissions by wisely scheduling clients.	Oper, NFunc, Biz	DoA, UC1
R3.3	Energy efficient ML architectures	The ROBUST-6G AI solutions (both centralized and decentralized) <u>should</u> integrate ML models that are energy efficient by design at inference time.	Oper, NFunc	DoA
R3.4	Trustworthy AI	The DFL framework developed in ROBUST-6G <u>should</u> be able to evaluate accountability, fairness, explainability and robustness in AI/ML models.	Oper, NFunc, Biz	DoA, UC1
R3.5	XAI practices	ROBUST-6G <u>must</u> employ robust Explainable AI (XAI) practices for threat detection, prediction, and mitigation, increasing transparency throughout these implementation processes.	Oper, NFunc	DoA, UC1
R3.6	Trustworthiness and robustness capabilities	The ROBUST-6G system <u>must</u> provide trustworthiness and robustness enhancement capabilities for AI-driven autonomous adaptations of 6G.	Oper, NFunc	DoA
R3.7	Attack prevention	The ROBUST-6G DFL framework <u>should</u> supply techniques to prevent attacks that attempt to infer an AI/ML model from spoofing learning messages flowing between federation nodes.	Tech, NFunc	DoA, UC1
R3.8	Secure communications	The ROBUST-6G system <u>should</u> make use of secure communication channels during the process of assessing the trustworthiness of the AI/ML models and the physical and sensing layers.	Tech, NFunc, Biz	DoA, UC1
R3.9	Data privacy	The ROBUST-6G system <u>should</u> ensure the privacy of data by implementing Differential Privacy (DP), Secure Multiparty Computation (SMC) or Homomorphic Encryption (HE) techniques in the FL framework.	Tech, NFunc, Biz	DoA
R3.10	Accountability in model lifecycle	The ROBUST-6G system <u>must</u> maintain detailed logs and audit trails for all stages of the AI/ML model lifecycle, including training, deployment, and updates, to ensure accountability.	Oper, NFunc, Biz	DoA
R3.11	Robustness testing	The ROBUST-6G system <u>must</u> incorporate extensive robustness testing, including adversarial attack simulations, to ensure model resilience against various threats such as poisoning or evasion attacks.	Tech, NFunc	DoA, UC1
R3.12	AI/ML tools and techniques	The ROBUST-6G system <u>shall</u> use AI/ML to enhance system security by facilitating predictive threat/anomaly detection.	Oper, Func	DoA
R3.13	E2E AI/ML driven	The ROBUST-6G system <u>should</u> introduce a Security-as-a-Service (SecaaS) based E2E AI/ML driven security framework for 6G.	Oper, Func, User	DoA
R3.14	Decentralized AI/ML	The ROBUST-6G system <u>must</u> generate and evaluate AI/ML models using a DFL framework for training shared models in a privacy-preserving manner by design.	Oper, Func	DoA, UC1
R3.15	Decentralized AI/ML	The fully DFL framework of ROBUST-6G <u>must</u> be agnostic of any application UC, applicable to any multiparty scenario where shared AI/ML models may be generated.	Oper, Func, Biz	DoA

R3.16	AI/ML models aggregation	The ROBUST-6G DFL framework <u>should</u> provide decentralized aggregation capabilities and local AI/ML model testing performed by multiple trusted entities, eliminating centralized aggregation processes that could cause bottlenecks and single point attacks.	Oper, NFunc, Biz	DoA, UC1
R3.17	APIs for AI/ML results	The AI/ML services associated with the DFL platform <u>should</u> provide a well-defined interface to obtain the results achieved by the AI/ML techniques used.	Oper, NFunc, Biz	DoA, UC1
R3.18	Model Performance Monitoring	The ROBUST-6G system <u>must</u> continuously monitor the performance of AI/ML models during the FL process to detect and mitigate any trustworthiness issues.	Oper, NFunc	DoA
R3.19	Fairness Assurance	The ROBUST-6G system <u>must</u> implement mechanisms to ensure fairness in the FL process. This includes identifying and mitigating biases in training data and model updates to ensure that AI/ML models do not unfairly benefit or harm any specific user group.	Oper, NFunc, Biz	DoA, UC1
R3.20	Trust and reputation management	The ROBUST-6G system <u>shall</u> provide a way to evaluate how inter-domain relationships behave using a reputation-based system approach.	Oper, NFunc	DoA, UC1
R3.21	User-centric controls	The ROBUST-6G system <u>should</u> provide user-centric controls, enabling users to manage their data and model preferences effectively.	Oper, NFunc	DoA

Finally, Table 3-6 lists a set of requirements to ensure integrated, hands-off security management for multi-tenant AI deployments in edge, fog and cloud environments, using APIs for security lifecycle management. Incorporated in this list of requirements are predictive threat mitigation, configurable incident response and secure data management with advanced monitoring tools, and semantic data analysis and visualisation with real-time and historical information.

Table 3-6: Zero-touch security management requirements in the ROBUST-6G system

Req. ID	Title	Requirement description	Type	Origin
Application Domain 4: ZERO-TOUCH SECURITY MANAGEMENT				
R4.1	Zero-touch security (ZTS) management	The ROBUST-6G platform <u>must</u> provide zero-touch integrated security management in multi-tenant distributed AI deployments.	Oper, Func	DoA
R4.2	Zero-Touch Security management	The ROBUST-6G platform <u>should</u> be able to manage security service requests including Security Policies or Security Service Level Agreements (SSLAs).	Oper, NFunc	DoA, ZTS
R4.3	External access APIs	The ROBUST-6G platform <u>must</u> define and provide a set of APIs specific for the lifecycle management of security services.	Oper, Func	ZTS
R4.4	Edge-to-Cloud management	The ROBUST-6G platform <u>should</u> be able to operate in multiple environments (edge, fog, cloud).	Oper, NFunc	DoA, ZTS
R4.5	Multiple Orchestrators	The ROBUST-6G platform <u>should</u> have different orchestrators (Security, Resources, Network) all connected.	Oper, Func	DoA, ZTS
R4.6	Multiple Closed-loops	The ROBUST-6G platform <u>must</u> support multiple closed loops and avoid conflicting configurations via a priority mechanism.	Oper, NFunc	DoA
R4.7	Zero-touch detection	The ROBUST-6G system <u>should</u> implement and use the zero-touch platform for threat detection and alarm generation.	Oper, Func	DoA, UC2
R4.8	Zero-touch Mitigation	The ROBUST-6G system <u>should</u> implement and use the zero-touch platform for deploying corrective/mitigation actions.	Oper, Func	DoA, UC2
R4.9	Threat prediction	The ROBUST-6G system <u>should</u> be able to predict incidents and impose corrective actions according to the predicted threat.	Oper, Func	DoA, UC2

R4.10	Security services heterogeneity	The ROBUST-6G platform <u>must</u> be able to manage heterogeneous security services with different requirements and capabilities.	Oper, Func	ZTS
R4.11	Point of investigation	The ROBUST-6G platform <u>should</u> provide a point of investigation for security experts to visualise security events and eventually the automated response is taken.	Oper, Func, User	ZTS
R4.12	Optimal mitigation strategy	The ROBUST-6G platform <u>should</u> provide optimal mitigation in terms of effectiveness and efficiency, considering security objectives and constraints.	Oper, NFunc	ZTS
R4.13	Incident response plans management	The ROBUST-6G platform <u>should</u> provide the possibility to define, modify and delete incident response plans in an incident response playbook.	Oper, Func, User	ZTS
R4.14	Threat Intelligence management	The ROBUST-6G platform <u>should</u> provide the ability to update and integrate from different sources, the Zero-Touch Security orchestrator threat intelligence.	Oper, NFunc	ZTS
R4.15	Efficient Resource Allocation	The ROBUST-6G platform <u>should</u> be aware of target environment resources and should be able to suggest a resource allocation strategy to targets while responding a threat.	Oper, NFunc	ZTS
R4.16	Data Collector	The ROBUST-6G monitoring platform <u>should</u> have multiple types of collectors to ensure flexible monitoring from several heterogeneous data sources at runtime.	Oper, Func	PMP
R4.17	Communication Bus	The ROBUST-6G monitoring platform <u>should</u> have a communication bus to forward the secure data parameters from collectors to the preprocessing modules.	Oper, Func	PMP
R4.18	Configuration Manager	The ROBUST-6G monitoring platform <u>should</u> have an entity in charge of interpreting the security requirements coming from the Security Orchestrator in order to deploy appropriate monitoring tools.	Oper, Func	ZTS, PMP
R4.19	Configuration Manager GUI or API	The ROBUST-6G monitoring platform <u>shall</u> provide a mechanism to enable external components or platform admin to interact with the Programmable Monitoring Platform (PMP).	Oper, Func	PMP
R4.20	Reconfiguration	The ROBUST-6G monitoring platform <u>should</u> support the on-demand configuration of their internal modules such as the Data Aggregation, Communication Bus, or Data Collection.	Oper, Func	PMP
R4.21	Maintenance	The ROBUST-6G <u>should</u> enable the Platform Admin to do the configuration of the PMP via a GUI.	Oper, User	PMP
R4.22	Access Control	The ROBUST-6G monitoring platform <u>should</u> have access control mechanisms to verify users trying to visualize data or add configurations have the privileges.	Oper, NFunc, User	PMP
R4.23	Short-term Data Storage	The ROBUST-6G monitoring platform <u>should</u> enable a database to store configuration parameters of monitoring tools or internal platform components.	Oper, Func	PMP
R4.24	Long-term Data Storage	The ROBUST-6G monitoring platform <u>should</u> enable a database to store collected and processed data in a scalable and secure manner. It also supports long-term storage for historical analysis.	Oper, NFunc	PMP
R4.25	Data Correlation and Feature Extraction	The ROBUST-6G monitoring platform <u>should</u> apply semantic techniques to understand the context, meaning, and significance of the monitored raw data, generating new features from the raw and correlated data that are more informative for prediction algorithms.	Oper, Func	PMP
R4.26	Data Exporter	The ROBUST-6G monitoring platform <u>shall</u> enable the export of data and report to external systems or for offline analysis.	Oper, Func	PMP

R4.27	Data Sources	The ROBUST-6G monitoring platform <u>should</u> provide a closed set of monitoring tools to ensure the proper acquisition of security params from network segments such as extreme-edge, edge, and cloud.	Oper, Func	PMP
R4.28	Data Storage GUI	The ROBUST-6G monitoring platform <u>should</u> support visualization capabilities for its long-term data storage in order to observe historical data, patterns, or analyse potential plots.	Oper, NFunc	PMP
R4.29	Historical Data Retrieval	The ROBUST-6G monitoring platform <u>shall</u> be capable of sharing its long-term data with external ROBUST-6G modules such as Data Management, Analysis Engines, or Alerting and Notification to perform more sophisticated activities.	Oper, Func	PMP, Data Fabric, ZTS
R4.30	Near Real-time Data Retrieval	The ROBUST-6G monitoring platform <u>shall</u> support external modules to consume real-time data to perform quick reactions or actions for their internal objectives.	Oper, Func	PMP, ZTS
R4.31	Alert Manager	The ROBUST-6G monitoring platform <u>should</u> generate alerts in case of unexpected behavior in aggregated information, stored in a Time Series Database of the PMP or the data pushed in the Data Fabric.	Oper, Func	ZTS, PMP

4 ROBUST-6G architecture

This section introduces the initial ROBUST-6G architecture, which is designed in alignment with the requirements defined in Section 3. The ROBUST-6G architecture, with integrated AI/ML techniques and security services exposure mechanisms, will enable E2E security in 6G networks. At the same time, the architecture design presented in this document is intended to support and enable a compatible interaction with the UC studies planned in the project.

In this direction, this section first introduces a high-level view of the proposed ROBUST-6G architecture, focusing on the structure from a broader aspect. The high-level architecture shows the system components by abstracting the detailed view on the services and interactions. This is then followed by the functional architecture of ROBUST-6G, which provides a higher resolution by depicting the internal functions, interfaces and services provided in the system. Moreover, each of the services in the architecture is introduced with a detailed inspection on how the components interwork in the common platform. Finally, a high-level deployment view of ROBUST-6G project contributions is presented in the envisioned 6G architecture [Hex24-D33], which is based on the principle of horizontal separation of the network functions from the underlying platform and overlying E2E management and exposure.

4.1 High-level ROBUST-6G architecture

The first high-level architecture of ROBUST-6G is illustrated in Figure 4-1. The main purpose of presenting this architectural view is to provide an overview of the technologies to be developed in the project and the objective of the system aiming to realize E2E, holistic security for the anticipated 6G networks. In this direction, the technologies being developed in the project and particular components included in each work package are shown in a modular manner.

On top of a common cloud infrastructure, the ROBUST-6G architecture envisions a 6G network deployment that provides not just connectivity services for the User Equipment (UE) and end devices but also value-added security services for the external consumers and verticals. Within the ROBUST-6G platform, monitoring data from the infrastructure (i.e., fault/performance measurements) is accomplished by the Programmable Monitoring Platform (PMP), which may help drive incident reports and alarms through other components that consume PMP outcomes. This entity continuously fetches the data from the underlying infrastructure and Network Functions (NFs) and then shares it with the Data Management Platform for internal access.

The main responsibility of the Data Management Platform is to enable distributed data management and implement relevant mechanisms for secure data access by the other entities. These entities are distributed in the management layer and network layer. While Physical Layer Security services are deployed in the Radio Access Network (RAN) domain, Security and AI Services implement a set of functions in the management layer to provide services for both internal and external use.

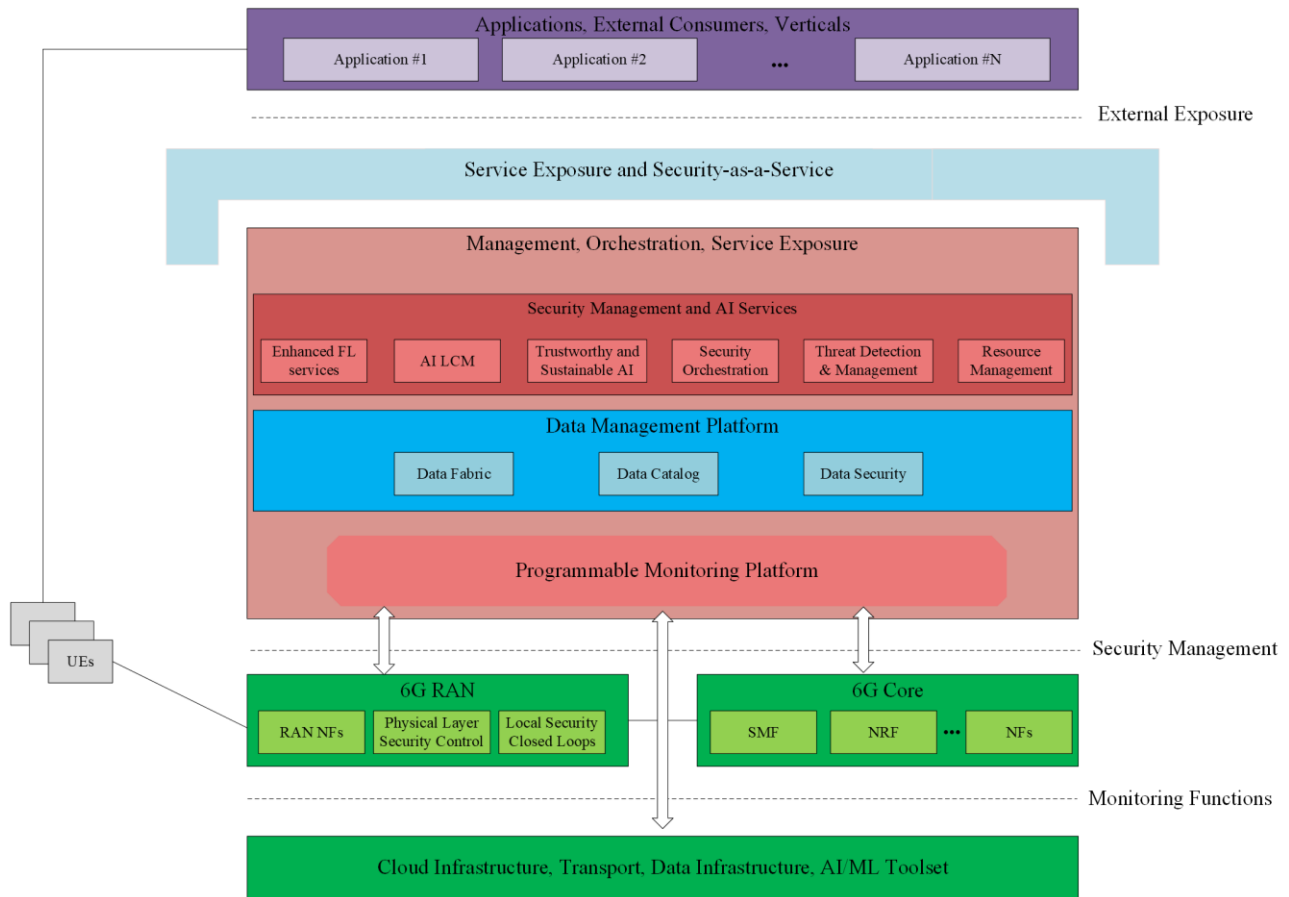


Figure 4-1: High-level architecture of ROBUST-6G project

Physical layer security will integrate AI-based mechanisms to process the signals and measurement/statistics data to detect anomalies and take proper actions when necessary. This function implements a local closed-loop mechanism to execute monitoring, analysis, and actuation processes. This closed-loop mechanism, with a broader scope, is also available in the management layer. With the vision of zero-touch management for secure 6G systems, multiple closed-loops are also part of the management layer for automated threat detection, prediction and mitigation during the security orchestration and resource management functionalities.

These AI-driven systems which will enable physical security and zero-touch security management are supported by the solution that has the role of providing trustworthy and sustainable AI/ML. With the accommodation of enhanced FL services and AI Lifecycle Management (LCM) capabilities, the physical layer security and zero-touch security management entities can be effectively managed in the network in a scalable manner.

On top of the management layer, we envision an exposure layer which will provide Security-as-a-Service (SecaaS) capability. By exposing the security services to the external consumers, ROBUST-6G platform will enable AI/ML driven SecaaS solutions, and so the external consumers can benefit from the internal features to indicate their requirements through the common exposure framework.

4.2 Functional architecture of ROBUST-6G

The high-level view of the initial ROBUST-6G architecture abstracts the interaction between the system components, details of the security services and the envisioned functionalities within. In this subsection, we introduce the functional architecture presenting a higher resolution.

As depicted in Figure 4-2, the proposed ROBUST-6G architecture is built on top of a common compute infrastructure which accommodates edge-cloud continuum. The NFs in RAN and Core domains are deployed over this compute infrastructure. Depending on the deployment model of the future 6G network, 6G RAN functions and 6G Core Network (CN) functions are distributed between far edge, near edge and central cloud systems. Besides, it should be noted that the deployment of virtualized NFs also depends on the requirements of the network services.

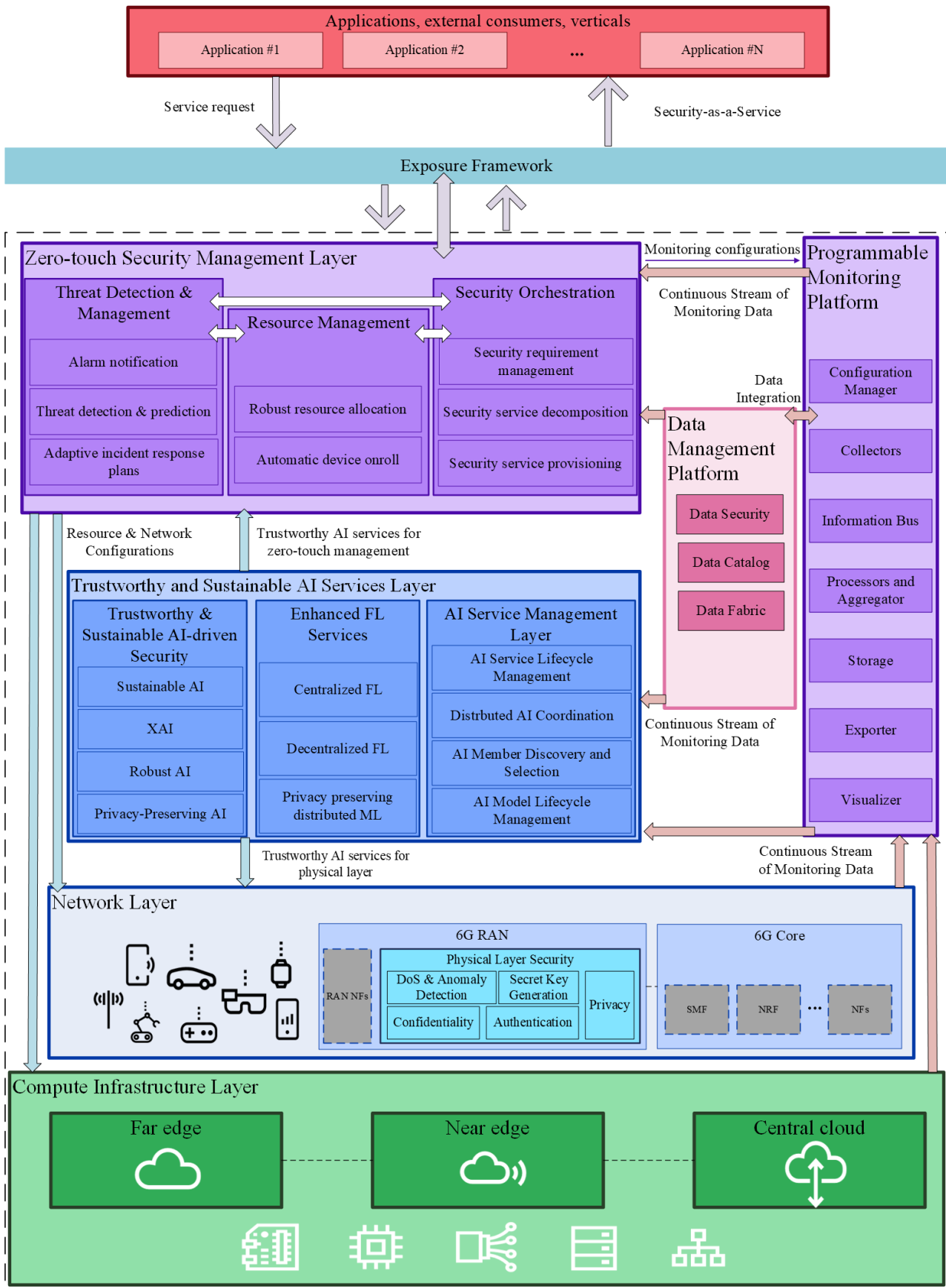


Figure 4-2: Functional architecture of ROBUST-6G project

One of the key components in the proposed ROBUST-6G platform is the Programmable Monitoring Platform, which is responsible for monitoring the underlying infrastructure (e.g., fault/performance measurements, incident reports, alarms) and sharing with the Data Management Platform. The data across different layers of the 6G system (e.g., network, application, and physical) are collected and aggregated in a unified framework.

In line with the 6G vision, ROBUST-6G supports three main aspects to realize autonomous security: (i) a zero-touch security management and orchestration mechanism powered by distributed and trustworthy AI; (ii) a common distributed data management platform in coordination with the pervasive monitoring; and (iii) exposure of security services where applications/customers can get easy access to exposed security services APIs.

4.4 ROBUST-6G security services in the architecture

The communication between the services and entities proposed to be offered on the ROBUST-6G platform should be examined in detail. Although different technologies are developed in a modular manner, these technologies should interwork in a compatible manner and consume the offered services accordingly to ensure the security of the entire system and to provide security services to the externals.

In this direction, this section discusses the objective of each technology and services to be provided in the common platform. Data Management Platform, Programmable Monitoring Platform, Physical Layer Security, Trustworthy and Sustainable AI Services and Zero-touch Security Management are the key technologies to be developed in the ROBUST-6G project. And they should coexist in this setting by consuming the services provided for internal use.

4.4.1 Data Management Platform

The Data Management Platform plays a central role in enabling and supervising the entire flow of data within the ROBUST-6G dataspace, with a strong emphasis on security and governance. This module is designed to handle the gathering, processing, and management of security-related data from multiple domains, including infrastructure, network, and services. Data is gathered through the Programmable Monitoring Platform and other potential data sources, enabling comprehensive monitoring across the system.

The Data Management Platform aligns with cutting-edge paradigms in distributed data management: Data Mesh and Data Fabric. These paradigms address the challenges of modern data management, particularly in environments with complex and rapidly changing data sources, like those anticipated in 6G networks.

- **Data Mesh** emphasizes a decentralized approach to data management. In this model, data is organized into domains, with each domain having dedicated ownership and stewardship responsibilities [CVH24]. This allows for:
 - **Domain Ownership:** Domain-specific teams manage the full lifecycle of data.
 - **Data-as-a-Product:** Data is treated as a valuable product, enhancing usability and quality.
 - **Self-Serve Data Infrastructure:** Domains can independently manage their data through shared infrastructure.
 - **Federated Governance:** Governance is distributed across domains to maintain compliance and consistency without central bottlenecks.
- **Data Fabric** complements the Data Mesh by realizing the self-serve data infrastructure that integrates data from diverse heterogeneous data sources, providing uniform access to data within the data spaces. Data Fabric builds upon a semantic layer introduces an integration layer that gathers data from various sources and transforms it to be consumed by applications such as AI/ML systems and analytics tools (see Figure 4-4). A notable feature of the Data Fabric is the ability to handle bidirectional data flows, depicted by the Reverse Extract-Transform-Load (ETL) block. This functionality allows systems to alternate between data consumers and sources based on the needs of the data process. For instance, while an AI/ML application may initially consume data to generate predictions, it can also serve as a source by sharing these predictions as new data [Gar24].

By combining the Data Mesh and Data Fabric paradigms, the Data Management Platform provides a robust, secure, and scalable framework for managing data in a distributed environment. This structure is essential for supporting the complexities of future 6G networks, ensuring effective governance, protection, and utilization of data across multiple domains. Through this innovative approach, the module supports the creation, exchange, and governance of data products, enhancing both security and functionality across the system.

The module features advanced discovery tools to identify assets that need protection and assess the security risks to which they are exposed. This is crucial in dynamic environments, such as future 6G networks, where assets and security threats are continually evolving.

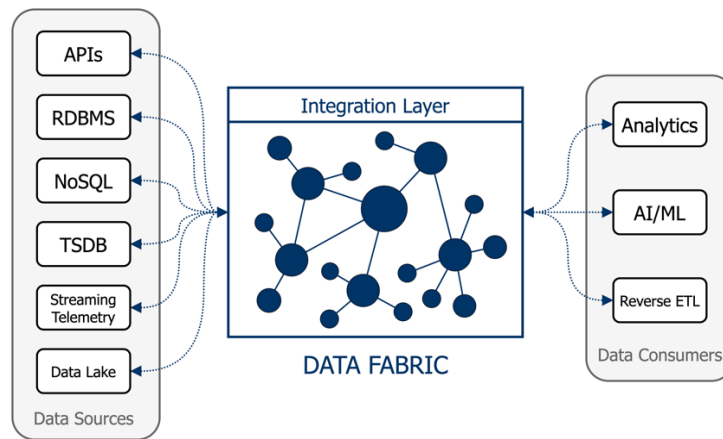


Figure 4-4: Data Fabric architecture

The Data Management Platform comprises two main building blocks:

- **Data Fabric:** Responsible for gathering, processing, and storing security-related data, the Data Fabric ensures that all monitored data is available to consumers via a unified, secure interface. It facilitates seamless data integration across diverse and heterogeneous sources such as relational databases, data streams, and APIs.
- **Data Governance:** This building block is composed of the following components:
 - **Data Catalog:** Provides a registry of the data products available within the data space, allowing users to find and access data according to predefined policies. The Data Catalog promotes data democratization by providing mechanisms for users to discover available data, while improving the quality, privacy, and trustworthiness of the catalogued data.
 - **Data Security:** Enforces robust access control policies that align with the security requirements specified by data product owners, determining precisely who can access specific types of data and under what conditions. By implementing stringent security protocols, it ensures full compliance with privacy, security, and regulatory standards.

Additionally, digital signatures are applied with verification conducted to confirm data origin and integrity. These signatures provide a trusted provenance, ensuring that data consumers can verify both the source and authenticity (non-repudiation) of the information. This added layer of security not only safeguards data integrity but also enhances confidence in data use by enabling traceability and accountability. Through this process, each data product within the system obtains an auditable history, which is critical for maintaining trust across the knowledge graph.

The Data Management Platform will be critical in ROBUST-6G, which implements scalable and dynamic data exchange mechanisms. It is designed to support AI/ML-based applications, such as threat detection systems, which rely on real-time data integration from various domains. The Data Fabric ensures efficient data handling across domains while maintaining data governance to protect sensitive data and manage access control.

4.4.2 Trustworthy and Sustainable AI Services

The Trustworthy and Sustainable AI Services layer integrates a number of interconnected components to provide distributed, secure, privacy-preserved, interpretable, and sustainable AI operations required for the development of 6G networks. This layer of the ROBUST-6G architecture is intended to ensure that AI systems are transparent, privacy-focused, robust, and sustainable, while also satisfying the security and performance requirements of 6G infrastructures. For example, XAI increases transparency by making AI models more interpretable, so ensuring that AI-driven security decisions are understood. This transparency is critical for building confidence in 6G services, which will be based on complicated, data-intensive procedures. At the same time, Sustainable AI attempts to limit the environmental effects of AI, which is an important factor for 6G networks that would require energy-efficient techniques to support a large number of connected devices. This is consistent with novel energy efficient learning approaches and the value of Distributed AI Coordination, which enables eco-friendly operations over vast, decentralized 6G networks where effective resource management is critical.

Privacy-preserving AI is essential in FL settings, especially within 6G networks, where sensitive data must be protected during collaborative training across distributed nodes. With 6G expected to support a massive number of devices and data-driven applications, robust AI is critical to safeguarding models against adversarial attacks, particularly in decentralized networks that are inherently more vulnerable. Enhanced FL Services address this challenge by enabling collaborative AI training in 6G. These services support both Centralized Federated Learning (CFL), where a central server aggregates models, and DFL, where nodes interact directly in a peer-to-peer manner. This flexibility enhances the scalability of 6G applications.

In both centralized and decentralized FL setups, both CFL and DFL, privacy-preserving techniques are crucial to securing data, especially given 6G's role in handling sensitive, mission-critical information. Privacy-preserving distributed ML further strengthens security by ensuring AI training across distributed environments complies with stringent privacy and security requirements while supporting the growing demand for AI-driven services.

The Trustworthy and Sustainable AI Services layer is important in managing AI operations within 6G networks. It ensures smooth operations and oversees the LCM of AI services, from model creation and training to deployment, adapting dynamically to the evolving needs of 6G applications. Distributed AI Coordination facilitates secure collaboration between nodes in both FL setups, maintaining data privacy despite the vast number of connected devices. AI Member Discovery and Selection mechanisms ensure that only trusted entities participate in distributed AI tasks, bolstering network security and trustworthiness. Finally, AI Model Lifecycle Management oversees the deployment, updates, and maintenance of AI models, ensuring they remain secure, effective, and resilient against emerging threats as 6G networks continue to grow.

Together, these elements create a robust and dependable framework for distributed AI services within the 6G ecosystem. By integrating privacy-preserving techniques, explainable AI models, and sustainable practices, this architecture meets the rigorous demands for security, privacy, transparency, and efficiency required by next-generation networks. It guarantees that AI systems within 6G are resilient, trustworthy, and sustainable, even in highly decentralized environments, supporting the future of AI-driven services central to 6G technologies.

4.4.2.1 *Explainable AI (XAI)*

XAI-Driven Enhancements for Federated Learning, Network Security, and Intrusion Detection in 6G Networks: Integrating XAI into FL frameworks offers transformative advancements in privacy, security, and model explainability, essential for next-generation wireless networks. By embedding interpretability and transparency into FL models, XAI enables a deeper understanding of decision-making processes, making it possible to identify vulnerabilities, mitigate attacks such as poisoning or inference threats, and optimize model behaviours. In federated architectures, where data privacy and robustness are critical, XAI bridges the gap between performance and trust by ensuring stakeholders can validate model decisions without compromising security. Implementing XAI in hierarchical or DFL systems further enhances scalability and distributes computational loads efficiently, supporting real-time adaptive decision-making across distributed systems. These contributions align directly with WP4 (“*Zero-Touch Management for secure 6G systems*”)’s objectives on adaptive threat detection and WP5 (“*AI/ML Enabled Physical Layer Security*”)’s focus on enhancing physical-layer security through privacy-preserving and explainable methods. The synergy between XAI and FL extends to ZSM and physical (PHY) layer applications in 6G networks.

FL facilitates privacy-preserving, distributed learning, enabling adaptive security solutions such as anomaly detection and interference mitigation. Coupled with XAI, these systems gain interpretability, allowing stakeholders to assess and trust automated security actions. For PHY-layer security, FL combined with XAI provides private and secure collaboration for detecting threats like signal interference or adversarial tampering, offering real-time, interpretable insights into network vulnerabilities. These efforts contribute to WP4 by improving automated and transparent security actions and to WP5 by addressing advanced PHY-layer threats in real-time.

Resilient Beamforming and Adversarial Attack Detection: Beamforming in mmWave communication systems is a key area where DL models enhance efficiency but face susceptibility to adversarial perturbations that threaten network stability. To address these challenges, we propose an XAI-based framework leveraging Shapley Additive Explanations (SHAP) values for adversarial attack detection and mitigation. Using explanation distillation, a “teacher” detector model trains a generalized “student” model to improve detection rates, especially for unseen attack scenarios. Recent studies [CCM21, ZMZ+22] have shown that DL models for MIMO-based communications are vulnerable to evasion attacks, degrading signal quality and throughput.

Current defences, including adversarial training and robust beamforming algorithms [KCC+23], often fail to generalize to new attack scenarios. The proposed XAI-based framework addresses these gaps, enabling real-time classification of threats and filtering them from benign received signals. This work aligns with WP4 Task 4.2 (“*Automatic Monitoring, Threat Detection, Alarm Generation*”), focusing on automated threat detection in ZSM, and WP5 Task 5.1 (“*Classification, Identification and Mitigation of Attacks at PHY*”), which targets mitigation of adversarial attacks in the physical layer. It ensures robust and secure beamforming operations in dMIMO environments.

XAI-Driven Intrusion Detection Systems: Incorporating XAI into IDSs enhances transparency and efficiency in ML-based threat detection. By using SHAP to analyse feature contributions across various ML models, our approach offers both global and local explanations for IDS decisions. This insight enables the development of robust clustering frameworks that optimize detection capabilities by prioritizing features and attack categories, reducing false positives while maintaining interpretability. Through rigorous data analysis, we observed similarities and differences in feature contributions and model behaviours, enabling the IDS to generate feature subsets for attack-specific detection, reducing model complexity without sacrificing performance. Cluster prioritization enhances targeted threat detection while maintaining explainability. The automated feedback loop aligns with WP4’s objectives on dynamic, zero-touch threat detection and incident prediction while optimizing computational resources. Furthermore, the transparency provided by XAI fosters trust in autonomous security mechanisms, directly addressing WP5’s goals of privacy-by-design and trust-building for context-aware and semantic security in 6G networks. This framework also supports WP5’s focus on integrating explainable, robust AI mechanisms into physical-layer security, ensuring timing and efficiency requirements in high-speed environments.

4.4.2.2 *Sustainable and scalable AI*

Because 6G devices process a tremendous amount of data to provide the users with connected intelligence and enhance the functionalities of the network control plane, distributed AI and ML must be scalable and energy-efficient in both the training and inference phases. The trustworthy and secure federated and fully decentralized training procedures developed must meet the energy requirement while ensuring an accurate learned global model. To reach this goal, the sustainable AI module provides client scheduling and robust aggregation algorithms via the optimization of dedicated KPIs, including measures of network (bandwidth, power) and local computing (CPU, GPU) resource consumption, and semantics-aware metrics, such as the linear and nonlinear version Age of Information (AoI). The latter metrics contribute to maximizing the usefulness of the information integrated into the model, hence reaching convergence faster and with less resource consumption.

By allowing the system to evaluate the significance of information exchange within the algorithm, semantics-aware metrics enable ISPs to avoid the need for fine-tuning training results to fit specific applications. This capability aligns seamlessly with the principles of zero-touch management. Semantics can also introduce context-adaptive features at the physical layer. For example, AoI-aware scheduling can enhance network efficiency by balancing resource utilization while ensuring QoS for critical users. Additionally, in scenarios where energy efficiency and reliability are paramount, semantics-aware metrics can dynamically adapt modulation, coding schemes, or error-correction mechanisms based on the significance of the data. This is particularly beneficial for power-constrained IoT devices, where critical updates, such as voltage control, can take priority over routine status checks.

Besides efficient decentralized training algorithms, the sustainable AI module provides the system with ML architectures that are sustainable by design at inference, which is the primary use of in-network AI. Specifically, spiking neural networks, RNNs processing event-based data closely inspired by the human brain, can reduce energy consumption by three orders of magnitude when running on dedicated neuromorphic hardware. Also, the sustainable AI module will be used by ZSM layer to perform energy-efficient continuous monitoring of automation tasks and threat detection during inference. Efficient and fast inference is also useful to PHY layer, where its signal processing can be enhanced by running real-time ML algorithms.

4.4.2.3 *Robust AI*

Prediction confidence metrics are integral to building robust and trustworthy AI systems, especially in complex and high-stakes environments like 6G networks. These metrics provide a quantitative measure of the reliability of model predictions, enabling systems to adapt dynamically to uncertainty, mitigate risks, and maintain consistent performance. By integrating confidence estimation into AI workflows, models can not only flag uncertain or anomalous predictions but also provide actionable insights that enhance decision-making, strengthen operational resilience, and build user trust.

Variational Autoencoders (VAEs) are leveraged for robust AI to derive a confidence metric based on latent space representations, a method particularly effective in applications like IDSs. This metric evaluates the trustworthiness of predictions by analysing their proximity to known data distributions in the latent space. This can serve as input to the IDS system of the architecture as a means to better detect and classify malicious activities/patterns in network traffic, thus reducing false positives and prioritizing high-certainty security predictions in the architecture.

Incorporating confidence metrics into XAI frameworks further improves the interpretability and resilience of these systems. XAI methods, such as those used in IDS or FL settings, can integrate confidence metrics to validate and refine explanations, ensuring they align with the underlying data distribution and model behaviour. For instance, in 6G applications like DFL for network optimization, confidence-aware XAI can provide transparency into model updates, facilitating trust and collaboration among nodes. This capability is particularly crucial in hierarchical or DFL architectures, where scalable and reliable decision-making processes must distribute computational loads efficiently without compromising privacy or accuracy.

Prediction confidence also plays a critical role in enhancing specific UCs across 6G networks. For instance, in beamforming models, confidence metrics can assist in detecting unseen adversarial attacks, ensuring resilient and adaptive communication strategies. Likewise, in a privacy-preserving FL approach, confidence-aware methods enhance model robustness and scalability by prioritizing high-certainty updates and optimizing resource usage across distributed nodes. Confidence-aware intrusion detection, combined with real-time XAI insights, ensures timely and effective responses to network anomalies. By embedding these mechanisms into the AI lifecycle, from training to deployment, we create systems that are not only robust and reliable but also adaptable to the evolving demands of next-generation network technologies. This synergy between prediction confidence and UC-specific applications highlights its central role in ensuring trustworthy AI for 6G. From improving anomaly detection in IDS to enabling scalable, energy-efficient learning in decentralized networks, confidence-aware mechanisms strengthen the resilience, transparency, and reliability of AI systems, thus supporting their seamless integration into critical 6G environments.

4.4.2.4 *Privacy-preserving AI*

In 6G networks, it is essential to ensure that sensitive data, which is normally used to train AI models, remains secure and confidential. As 6G networks are anticipated to incorporate advanced AI technologies for tasks like network optimization, predictive analytics, and user behaviour analysis, ensuring privacy while leveraging AI's capabilities is crucial. Privacy-preserving techniques, such as FL, Homomorphic Encryption (HE), and Differential Privacy (DP), can be utilized to enhance AI privacy within the 6G network. FL enables distributed training, where data remains local on devices, and only model updates are shared, protecting user data from exposure. By integrating FL, 6G networks can offer privacy-preserving solutions enabling secure and efficient AI-driven functionalities. Even more, to enhance the privacy of FL while model updates are shared with the server, HE can be used to encrypt model updates before being sent to another element for aggregation. HE allows computations (i.e., aggregation) to be performed on encrypted data at server side without needing to decrypt them. This ensures that sensitive data never leaves the device and remains secure throughout the FL process. FL service can be used in different areas of interest in the network; it can be used by network layer and Zero-touch Security Management Layer as illustrated in Figure 4-2. Within such an area, the client nodes, such as sensors, BSs, or edge devices in case of RAN or NFs in the CN, may collaborate to train a ML model without sharing data. A client node in these areas has access to regional data and training capabilities and sends updated model weights to the server. A server is in charge of model aggregation and coordination of the FL process. The Distributed AI Coordination component in the Trustworthy and Sustainable AI Services layer (Figure 4-2) will be responsible to coordinate the FL process.

In addition, the AI functionality which will be used in 6G network, is vulnerable to privacy attacks such as model inversion attacks and membership inference attacks. This is an important concern specially when AI systems process sensitive user data in the network. Privacy techniques such as DP, Secure Multi-Party Computation (SMPC), Anonymization and Data Masking can be used to prevent privacy leakage. Thus, the AI functions which are used by different layers in 6G network can request Privacy-preserving AI component for privacy services. The privacy services will be provided with respect to the UC, the requirements, and capability of the devices.

4.4.2.5 *Enhanced FL Services*

Traditional centralized data processing approaches are inadequate for decentralized and dynamic environments like the ones expected for 6G, due to scalability, latency, and privacy limitations. Enhanced FL Services have

emerged as a pivotal solution within 6G networks, enabling collaborative ML processes across distributed devices without the need to centralize sensitive information. These services leverage the computational capabilities of edge devices and network nodes, facilitating real-time analytics and decision-making directly at the edge of the network. This approach reduces communication overhead and enhances data security by keeping sensitive information localized.

Centralized FL

CFL enables multiple clients—such as user devices, edge nodes, or other computational entities within a network—to collaboratively train a shared global model under the coordination of a central server or aggregator. In this paradigm, each client uses its local data to train a local model and then periodically transmits updates, typically in the form of model parameters or gradients, to the central aggregator. The central aggregator is responsible for collecting these updates from all participating clients, integrating them to refine the global model, and then redistributing the updated global model back to the clients for further local training. In the context of 6G networks, CFL can be particularly effective when the network architecture and model design allow for centralized aggregation without incurring significant communication overhead or latency. Specific nodes within the network such as edge servers or cloud-based central units that have sufficient computational resources and network connectivity can serve as central aggregators. These nodes facilitate the aggregation process by efficiently managing the collection and dissemination of model updates. However, CFL presents several challenges, particularly in 6G networks with massive device connectivity and stringent latency requirements. The central aggregator can become a bottleneck as the number of clients increases, potentially leading to increased latency and reduced scalability. Moreover, the central aggregator represents a single point of failure; if it becomes unavailable or compromised, the entire training process can be disrupted. This centralization can also raise privacy concerns, as the aggregator may have visibility into the model updates, which could potentially be exploited to infer sensitive information about the client's local data.

Decentralized FL

DFL is an advanced collaborative ML paradigm that eliminates the need for a central server or aggregator. In contrast to traditional CFL, DFL allows clients such as mobile devices, edge nodes, or other network entities to communicate directly with each other to share and update AI/ML models. This decentralized approach leverages the distributed nature of modern networks, particularly in 6G environments, enhancing scalability, robustness, and privacy. In DFL, the learning process is entirely distributed among the participating clients. Each client maintains its local model and uses its private data to train this model. The clients periodically exchange model updates with their peers, collaboratively working towards a consensus on the global model without relying on a central coordinator.

Initially, each client independently initializes its local model, which could be a random initialization or based on a commonly agreed upon starting point. The clients then perform local training using their private datasets, updating the model parameters to reflect patterns and insights derived from their local data. This local training phase is crucial, as it allows clients to learn from their data without exposing it to others, preserving privacy. Following local training, clients engage in peer-to-peer communication to exchange model updates. Peer selection is a critical component of this process. Clients may select peers based on various criteria, such as network topology to optimize connectivity, trust levels to ensure security or even random selection to enhance robustness. Effective peer discovery mechanisms are essential to identify suitable peers for model exchange, balancing the need for diversity in the updates received with practical network efficiency considerations. Once peers are selected, clients securely exchange their model updates. Secure communication protocols are employed to protect the confidentiality and integrity of the model updates during transmission. This is particularly important to safeguard against potential adversarial attacks.

Upon receiving model updates from their peers, each client performs local aggregation. This aggregation step integrates the received updates with the client's model parameters to create an updated local model. Aggregation methods can vary in complexity. Simple approaches might involve averaging the model parameters, while more sophisticated methods might weigh up updates based on the reliability of the source, the relevance to the client's data, or the statistical properties of the updates. For example, clients might assign higher weights to updates from peers with similar data distributions or those with a history of providing high-quality updates. The updated local model is then used for further training, which repeats in subsequent iterations. This iterative process progressively allows the models across the network to converge toward a consensus. Over time, despite the lack of a central coordinator, the collaborative learning process enables clients to build more generalizable and robust models than those trained solely on local data.

Privacy-preserving distributed ML

Privacy-preserving distributed ML is a cornerstone of deploying collaborative AI models in 6G networks, where protecting sensitive data and resilience against adversarial attacks are paramount. As devices participate in FL processes, they exchange information that, if not properly secured, could lead to privacy breaches or be exploited by malicious actors. To address these challenges, advanced techniques such as DP and secure aggregation algorithms are employed to enhance both the confidentiality and integrity of the learning process. DP provides quantifiable privacy guarantees for individuals within a dataset. In the context of distributed ML, DP ensures that the inclusion or exclusion of a single client's data has a negligible impact on the output of the learning algorithm. This property makes it statistically improbable for an adversary to infer sensitive information about any individual participant based solely on the model updates or the final model parameters. Implementing DP in FL involves adding carefully calibrated noise to the model updates before sharing them. Each client perturbs its local model parameters or gradients with random noise drawn from a specific probability distribution, typically a Gaussian or Laplace distribution. The amount of noise is determined by the desired level of privacy, quantified by the privacy loss parameter, often denoted as epsilon (ϵ). A smaller epsilon corresponds to stronger privacy but may impact model accuracy due to the introduced noise. The key steps in applying DP in distributed learning are:

- **Local Noise Addition:** Each client adds random noise to its model updates before transmission. This noise masks the contributions of individual data points, ensuring that sensitive information cannot be retrieved from the updates.
- **Privacy Budget Management:** The cumulative privacy loss over multiple training rounds is tracked using a privacy accounting mechanism. Clients manage their privacy budgets to balance model utility and privacy protection trade-offs.
- **Aggregation of Noisy Updates:** The aggregator (or peers in decentralized settings) combines the noisy updates. Due to the properties of DP, the aggregated model retains useful patterns from the data while preserving individual privacy.

Secure Aggregation Algorithms are designed to protect the integrity of the aggregation process in distributed learning, particularly against adversarial attacks that aim to disrupt or manipulate the global model. Unlike traditional aggregation methods like Federated Averaging (FedAvg), which may be susceptible to malicious clients injecting false or corrupted updates, secure aggregation algorithms incorporate mechanisms to detect and mitigate such threats. These algorithms focus on aggregation settings and strategies that enhance robustness against attacks, improving standard methods by incorporating statistical and algorithmic safeguards. In this sense, they account for malicious or faulty clients that may send arbitrary or harmful updates. The aggregator employs statistical techniques to filter out anomalous updates. For example, instead of computing a simple average, the aggregator might use coordinate-wise median or trimmed mean calculations, which are less sensitive to extreme values introduced by adversarial clients. In other perspectives, algorithms could select updates that are most representative of the majority, effectively isolating outliers that may result from malicious activity. These methods involve computing the distances between updates and selecting those that are closest to the consensus of the group. Combining DP with robust secure aggregation provides a comprehensive defense strategy. DP protects individual data privacy by obscuring specific contributions, while robust aggregation methods safeguard the overall model integrity against malicious attempts to disrupt the learning process. This dual approach is particularly effective in large-scale networks like 6G, where the diversity of devices and potential for adversarial behavior is significant.

4.4.2.6 AI Service Management Layer

Today's networks mostly provide communication capabilities, but 6G is expected to transform the network into a powerfully distributed AI platform. The vision for 6G demands for the widespread presence of AI applications, and this requires the evolution of mobile network architecture into a platform that supports such applications. Thus, future network architecture should be designed to support AI applications and services. The network is expected to be responsible for orchestrating, managing, scheduling, and exposing AI-related network services. This exposure of AI-related services can be enabled by the AI-as-a-Service (AIaaS) concept.

AIaaS consists of enablers and APIs offering AI functionalities to internal network/application functions, or external consumers. AIaaS provides AI services and functions such as analytics, prediction, classification, eliminating the need for users to build and maintain their own AI infrastructure. The AI services are supplied by pre-built AI models that are accessible via APIs. Possible examples for the exposed AI services can be a

network analytics service for different KPIs or security analytics service for security incident prediction or detection.

External and/or internal consumers can access the models that have been deployed in the AI agents within the AI services layer. Inference requests initiated by the consumers are handled by the AI agents and the outputs are in the form of predictions or reports for the customers in an as-a-service fashion. In such a scenario, deployed models might be vulnerable to several types of attacks, such as model evasion and model inversion, in which adversaries attempt to influence the decision of the target model by carefully constructed perturbations to the input. These risks necessitate constant monitoring of the inference outputs provided by AI agents and activating retraining procedures when necessary.

4.4.3 Zero-touch Security Management

Zero-touch security management encompasses many different and complex concepts. First of all, it includes SecaaS, which aims at the flexible implementation of scalable security solutions. These cybersecurity solutions are made transparent to the consumer in terms of implementation and infrastructure overhead. This reduces the need for consumers to be experts in technical security details, as they can simply request a high-level service and let the system handle it in the background. In addition, the consumer gets an easy-to-use solution that is, in most cases, continuously updated to the latest version. The Security as a Service functionality defined before is just an example of the whole potentiality of the zero-touch security management module. Other functionalities like monitoring, threat detection, threat prediction, threat mitigation, and resource orchestration are designed according to the requirements defined in Section 3, Table 3-6, reporting the zero-touch security management application domain.

The complexity of problems of these dimensions may require the adoption of standards and abstractions such as the closed-loop. The closed loop is a simple way of representing the sequence of tasks that follow during the application of a particular service. The concept of the closed loop has been already presented in Section 2.2, and more in detail in Deliverable 4.1 but to briefly remind, it is composed of the following four functions: monitoring, analysis, decision, and action.

The monitoring step may be implemented by one or more components, and it is responsible for collecting data from multiple data sources. The first step regarding the collection and classification of data from different data sources is extensively discussed in the previous Sections 4.4.1 and 4.4.2 through the use of the Data Management Platform and the Programmable Monitoring Platform.

The next two steps lead to the analysis and decision steps. They consist of the study of all collected data for the extraction of valuable information such as the presence of a threat. Later, the decision step defines which could be the most suitable plan to execute to mitigate the previously detected anomaly. From an alternative perspective, the analysis and decision-making phases of the closed loop can benefit from the integration of AI/ML models, which enhance the accuracy and efficiency of identifying the most appropriate actions. With the rapid escalation in the volume and complexity of cyber threats, traditional reactive approaches that depend on human intervention are increasingly ineffective. By adopting a modern zero-touch architecture coupled with DL models, organizations can not only accelerate threat detection and mitigation but also minimize human error and response delays, ultimately reducing the impact of attacks and limiting downtime.

Finally, the last step of a closed loop is the action, which translates into the physical execution of a decided action to update the state of the system to a new healthy one. To automate all of the steps, the concept of the security orchestrator plays a crucial role. In fact, an orchestrator minimises manual intervention, thereby reducing costs and human error, by abstracting very complex combinations of execution environments, themselves often composed with more dedicated orchestrators. A security orchestrator has a northern interface for consumers who need to address security services requirements towards a service provider. Once both parties agree on the service they wish to achieve, the orchestrator's role is to decompose and translate the requirements to a combination of requests towards orchestrators specialized for the targeted environments, using a southern interface such as the Network Orchestrator or the Resource Orchestrator. For example, these requests can contain the expression of the security requirements of a consumer for a core network environment, a cloud environment, or an edge environment. Then the goal of the security orchestrator is to ensure that the agreed conditions are met and maintained throughout time and in an E2E way, in all the execution environments, by gathering information from the underlying orchestrators. It means that if, when monitoring certain conditions, the analysis detects the presence of a threat to the consumer's services, a specific action should be decided and scheduled to resolve the issue in all the environments concerned by that threat. The actions to be computed and taken by the security orchestrator are policy-based. A security policy is the product

of the agreement between the service consumer, who wants to deploy its service with security, and a service provider, who owns a catalogue of security services to implement the policy. The security policy is applied over a topology of services which abstracts the environments where they are deployed. Through this security policy, a security orchestrator observes the state of the services, and alerts or remediates to threats and modifications, when a violation of the policy is detected.

4.4.3.1 Zero-touch Security Platform: a functional architecture

One of the key aspects characterizing the definition of the Zero-touch Security Platform architecture is the separation of concerns. The functional architecture is indeed built by several macro-services covering specific roles through a set of dedicated functionalities, insisting on any segment of the 5G/6G mobile network. Figure 4-5 shows the set of Macro-Services and their connections.

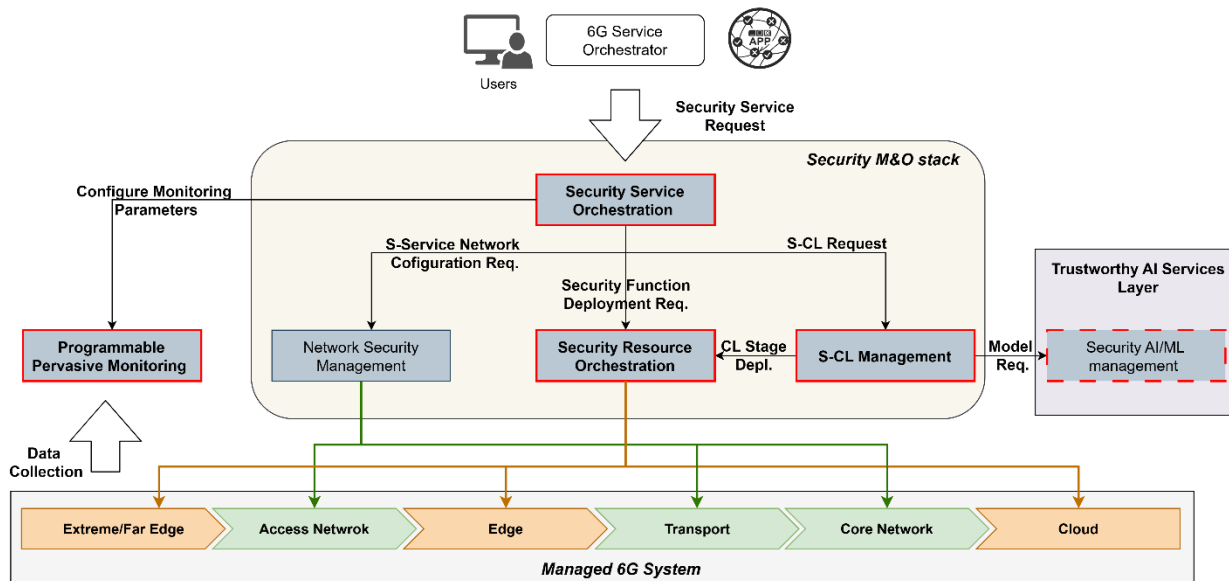


Figure 4-5: ROBUST-6G Zero-touch security platform functional architecture

The consumers of the platform security capabilities can be human users (e.g., verticals, sysadmins, etc.), third party applications and even existing orchestrators (e.g., 6G service orchestrators belonging to the Telco Provider). In this last case, the 6G orchestrator may request the provisioning of a security service to the zero-touch security platform as a side element of a vertical mobile service it is orchestrating. For example, a vertical may request the provisioning of a video stream service, specifying certain security constraints: the 6G orchestrator will request to the security platform a security service that addresses the security requirements, in a transparent manner from the vertical point of view, enabling the concept of SecaaS.

The exposure of the security capabilities is realized through a specific exposure layer (not shown in Figure 4-5) whose characteristics mainly depend on the level of integration with the existing 6G MNO infrastructure.

Security Service Orchestration (S-SO). Represents the core service for the security orchestration and the entry point for requesting Security Services (S-Services). Its internal logic encompasses functionalities for service parsing, validation, and decomposition, based on specific information models aimed at providing a uniform definition of S-Services. S-Service Orchestration also includes functionalities for maintaining the status of provisioned S-Services and the related remediation plans, executed every time an anomaly, targeted by the S-Service, is detected and/or predicted.

Security Resource Orchestration (S-RO). Performs deployment and configuration of the Security Applications (S-Applications) parts of the S-Service in the target environment i.e., Cloud, Edge, Far/Extreme Edge as shown in Figure 4-5, by exploiting existing cloud managers of those segments e.g., Kubernetes [K8s24], K3s [K3s24], OpenStack [OS24], etc. S-Application can be well-known security tools (e.g., Snort [Snort24]) and/or brand-new SW designed and developed from scratch for the purpose of the project. This would include also those applications in charge of implementing the different stages of the Security Closed-Loop (S-CL) in particular Decision and Execution: the Monitoring stage is realized through a dedicated programmable pervasive monitoring service (described below), while the Analysis stage can be orchestrated by either the S-RO or the Security AI/ML management service, as described below. S-RO also implements algorithms and techniques for robust resource allocation.

Network Security Management (NSMgmt). Applies specific network configuration to guarantee security on RAN, Transport and CNs, according to the security constraints specified by the S-Service. As per S-RO, relies on existing network controllers for each segment, e.g., Software-Defined Networking (SDN) controllers for Radio and Transport and CN control plane.

Security Closed-Loop Management (S-CLMgmt). This Macro-Service covers two important aspects characterizing the management of S-CLs, following the work published by ETSI ZSM in ZSM-009 [ZSM009-1]: i) S-CL Governance i.e., interfaces and logic for S-CL Orchestration, stage configuration, runtime interaction, and ii) S-CL Coordination i.e., interfaces and logic for frictionless co-existence of multiple closed-loops.

Programmable Pervasive Monitoring. Provides functionalities to select and collect data from different data sources suitable for the monitoring a target environment a timely detect/predict anomaly. The most important ones are the real-time data exposure, for immediate detection and/or AI/ML inference, and the historical data exposure, useful for AI/ML training and offline analysis. This Macro-Service is highly pervasive, capable of monitoring parameters at different layers and segments of the 6G systems.

Security AI/ML Management. Provides functionalities for secure AI/ML model training, selection of models and datasets, and possibly, AI/ML model deployment in the target environment. The model selected is part of the S-CL associated with a given security service, mainly covering the role of the Analysis stage of the loop, although it can be employed also for decision purposes. This Macro-Service is an abstraction that groups all the functionalities exposed by the AI Service Layer in the Trustworthy AI Services Layer shown in Figure 4-2, i.e., AI Service Lifecycle management, Distributed AI Coordination, AI Member Discovery and Selection, AI Model Lifecycle Management.

4.4.4 Physical Layer Security

Physical Layer Security is a key component of the ROBUST-6G project and is expected to play an important role in complementing cryptographic security in future cellular systems. To this end, a specific module is expected to operate on 6G RAN where physical layer signals are collected and where we can have a faster response to threats at the physical layer. The module will include several components that are related to the detection and mitigation of specific threats. In particular, the module will include the following components:

- **DoS and anomaly detection:** This model aims at detecting DoS attacks and in general anomalies. It will be based on the detection of jamming attacks and other DoS attacks that operate at the physical layer and other anomalies in the signals can be detected. This block may include ML parts for the automatic analysis of the signals and classification techniques to detect anomalies. It may leverage the RF features of devices as an input to detect attack attempts.
- **Secret key generation:** Key generation is an important security primitive for cryptography, and in Physical Layer Security keys can be obtained from the signals exchanged on air among devices. This module will have a strong interaction with the capabilities of the propagation environment available in the cell. For example, the availability of RISs and multiple antennas (MIMO) will be exploited to generate the key. Some interaction of local instances in different BSs can be envisioned when using dMIMO solutions operating over different BSs. Use of ML solutions for information extraction and information reconciliation will be used (possibly distributed in a federated approach) to learn the statistics of the channel and the most useful channel features to be used for randomness extraction. The ML model can be shared with other BSs, especially when they cover the same environment with overlapping of coverage.
- **Privacy:** Privacy-preserving technologies are important in all layers where sensitive data are stored or processed. Therefore, this module will integrate distributed privacy-enhancing mechanisms, that are relevant to establishing distributed trust, i.e., trust that is not anchored in a central trusted authority, to avoid some problems and concerns such as single points of failure and risks of data violations and privacy compromises. For this purpose, FL will be used to guarantee a fully distributed framework.
- **Confidentiality:** Confidentiality solutions have been the first to be considered in the physical-layer security context, with the study and design of wiretap coding solutions. This component will include such solutions in the architecture and will exploit specific new technologies that are under consideration for 6G networks, such as highly directive links, e.g., fronthaul/backhaul using a massive number of antennas in mmWave, RISs, and dMIMO and other controllable devices (such as drones operating as relays or BSs) are available. When using these technologies for confidentiality, ML solutions will be adopted to learn how to configure the devices and what signals they should transmit

to improve the confidentiality of transmissions. Such models will greatly depend on the local wireless propagation characteristics but can also benefit from common training among different BSs that cover partially overlapping environments.

- **Authentication:** Authentication is among the main security mechanisms that can greatly benefit from the elaboration of physical-layer signals, as several features of these signals uniquely identify devices. This capability is also critical for identifying fake BSs and spoofing attackers, as adversaries can deploy rogue transmitters to mimic legitimate ones, posing significant security risks. Indeed, two main approaches can be followed for authentication: either we use the impairments added by the transmitter on the transmitter data (in what is known as device fingerprinting) or we use the characteristics over which signals are transmitted to identify the location of the transmitter and, under time-unvarying conditions, its identify (in what is known as channel-based authentication). In both cases, the identification of the users is affected by noise and interference that degrade the estimation of the device fingerprint or channel characteristics at the receiver. Moreover, time variations should also be taken into consideration when deciding on the authenticity of the received signals. To this end, ML techniques can be useful and again we have models that can be trained and used either locally or in part shared among multiple BSs. Therefore, the authentication component will have interactions with other analogous components in other BSs and models can be exchanged and trained in a distributed / decentralized fashion in the network. Also, to consider time variations in channel characteristics, the autoregressive model approach might be utilized for system with low mobility.

All these components may use and share ML models to perform detection and mitigation. The training of such models get support from the envisioned Trustworthy and Sustainable AI services. This part gets the necessary data from the data management platform, helps to train and validate the models centrally in a trustworthy way and deploys it to the RAN domain. Therefore, the security orchestrator will organize the data flow and the decentralized/distributed learning, as well as the distribution of the models to RANs.

Furthermore, to give some examples of how these components impact the other components in the architecture, we provide the following three examples:

1. By combining robust and fast authentication using the Angle of Arrival (AoA) with fast secret key generation using Long Short-Term Memory (LSTM) networks, fast authentication and key agreement protocols can be provided. These are instrumental specifically for the UC2, scenario 3.
2. Working on proving that the AoA in digital MIMO arrays is a robust authentication feature, i.e., is not prone to impersonation type of attacks, allows to use it as a trusted feature to identify malicious behaviour in dynamic systems, e.g., vehicle to vehicle networks as was demonstrated in our works so far. This in the future can feed the update of reputation models used in WP3 (“*Trustworthy and Sustainable AI/ML for 6G Security*”).
3. Robust location authentication (e.g., using AoA) can be used in zero touch automaton for automatic device enrolment, which is examined in WP4.

The above advanced are currently in various stages of advancement and are expected to be delivered in their totality before M24.

4.4.5 Use Case Interactions

Use Case 1 (UC1) is thoroughly detailed in Section 2.1. Below, a global sequence diagram illustrates the main interactions between the various components involved in UC1 and the ROBUST-6G architecture (see Figure 4-6). The sequence diagram illustrates a process where a consumer, acting as the end-user, initiates a DFL service to enhance the performance and security of distributed AI applications. The primary objective is to leverage advanced AI techniques while ensuring data privacy, robustness against attacks, and model explainability.

Initially, the consumer requests the DFL process from the Enhanced FL Service. This service then communicates with the Distributed AI Coordination component to begin the DFL process which is part of the initialization phase. The coordination component deploys the necessary algorithms across the network through the DFL framework. This deployment allows for the training of local models on distributed devices or nodes, facilitating collaborative learning without centralized data aggregation.

As the local models are trained, trustworthy AI-driven security techniques are applied to ensure the integrity and reliability of the learning process. Privacy-Preserving AI solutions implement confidential computing methods on the data to be included in the framework, safeguarding sensitive information during training. The

Robust AI approach checks the performance of the trained models against potential adversarial attacks and ensures resilience against poisoning attacks that could compromise model integrity. Additionally, the XAI component utilizes methods like SHAP and Local Interpretable Model-Agnostic Explanations (LIME) as post-hoc explanations to provide insights into the model's decisions, enhancing transparency and trustworthiness.

After the models have been trained and secured, the framework shares the model updates across the network. This collaborative sharing enhances the overall model performance by aggregating knowledge from various local models while maintaining data privacy. The consumer then provides feedback on the model's performance to the Network Administrator. Based on this feedback, the administrator deploys a readjusted model across the network through the framework, ensuring that the models continue to meet the user's needs and adapt to new data or requirements.

The main objectives of this process are to facilitate a secure and efficient decentralized learning environment that respects user privacy, maintains robustness against security threats, and offers explainable insights into AI decisions. The interactions between the consumer, AI components, and network administrator work cohesively to achieve these goals, resulting in a trustworthy and adaptive AI system that aligns with the consumer's expectations and the dynamic nature of distributed networks.

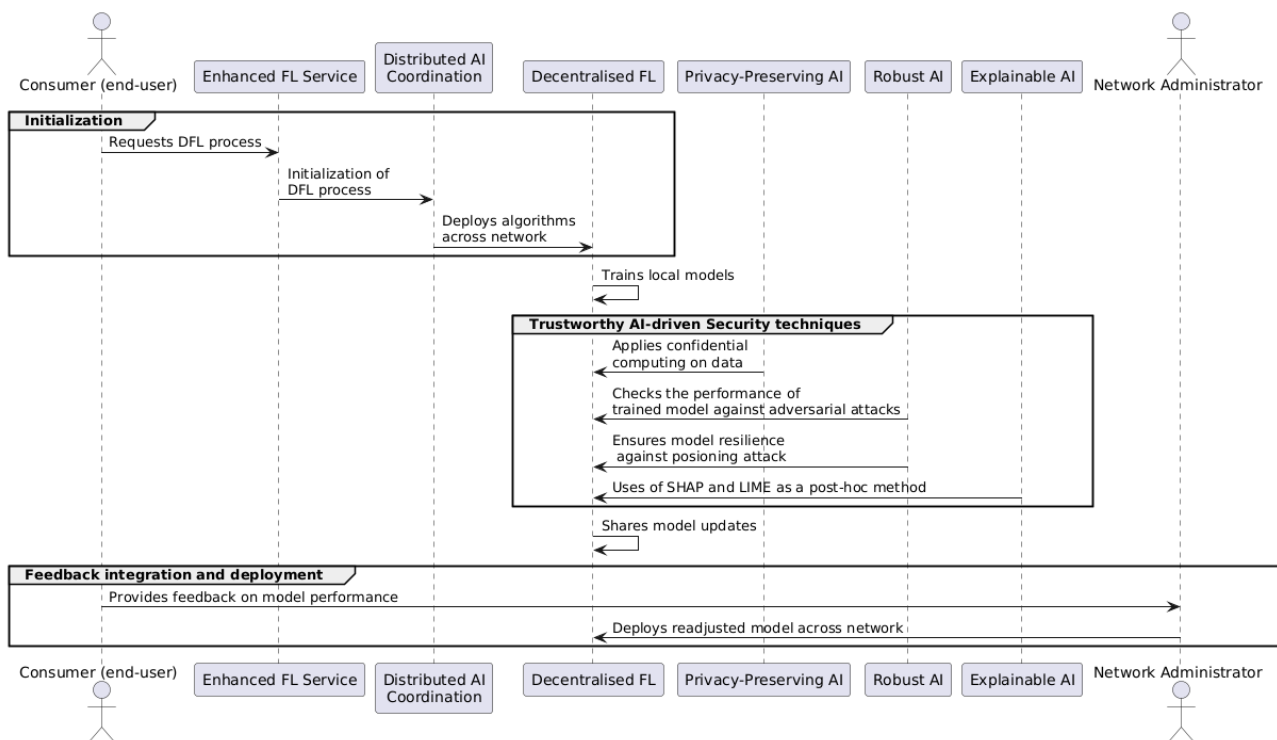


Figure 4-6: UC1 Scenario 1 interactions

Use Case 2 (UC2) has been deeply described in Section 2.2 as well as in Section 4.4.3 with a focus on technical details and presenting possible components implementing the proposed functionalities. A high-level sequence diagram mapping these functional elements' interaction is reported in Figure 4-7.

As depicted in the diagram, the workflow can be divided into two distinct phases. The first phase, usually identified as Pro-Active Security Orchestration, is responsible for the initial configuration that provides the necessary parameters to correctly handle the services and the associated security constraints. This fact means that once an external consumer executes a service request, the Zero-touch Security Orchestration interacts with both the Data Monitoring & Management module and, on the other hand, the Trustworthy and Sustainable AI Services module.

The interaction with the Data Module triggers the data collection process. The other interaction with the AI Module is required for getting suitable ML models capable of process and analyse the collected data. It is important to emphasise that these preliminary steps are essential for the proper configuration of a Security Application in the target environment, ensuring the execution of the close-loop functions during run-time operations. The second phase also known as Reactive/Predictive Security Orchestration happens at runtime.

During this phase, the security application continuously receives the collected and structured data from the environment and uses the preconfigured ML models for data analyses. If the analysis identifies the presence (or predicts in case of the future) of a threat two alternative scenarios come into play.

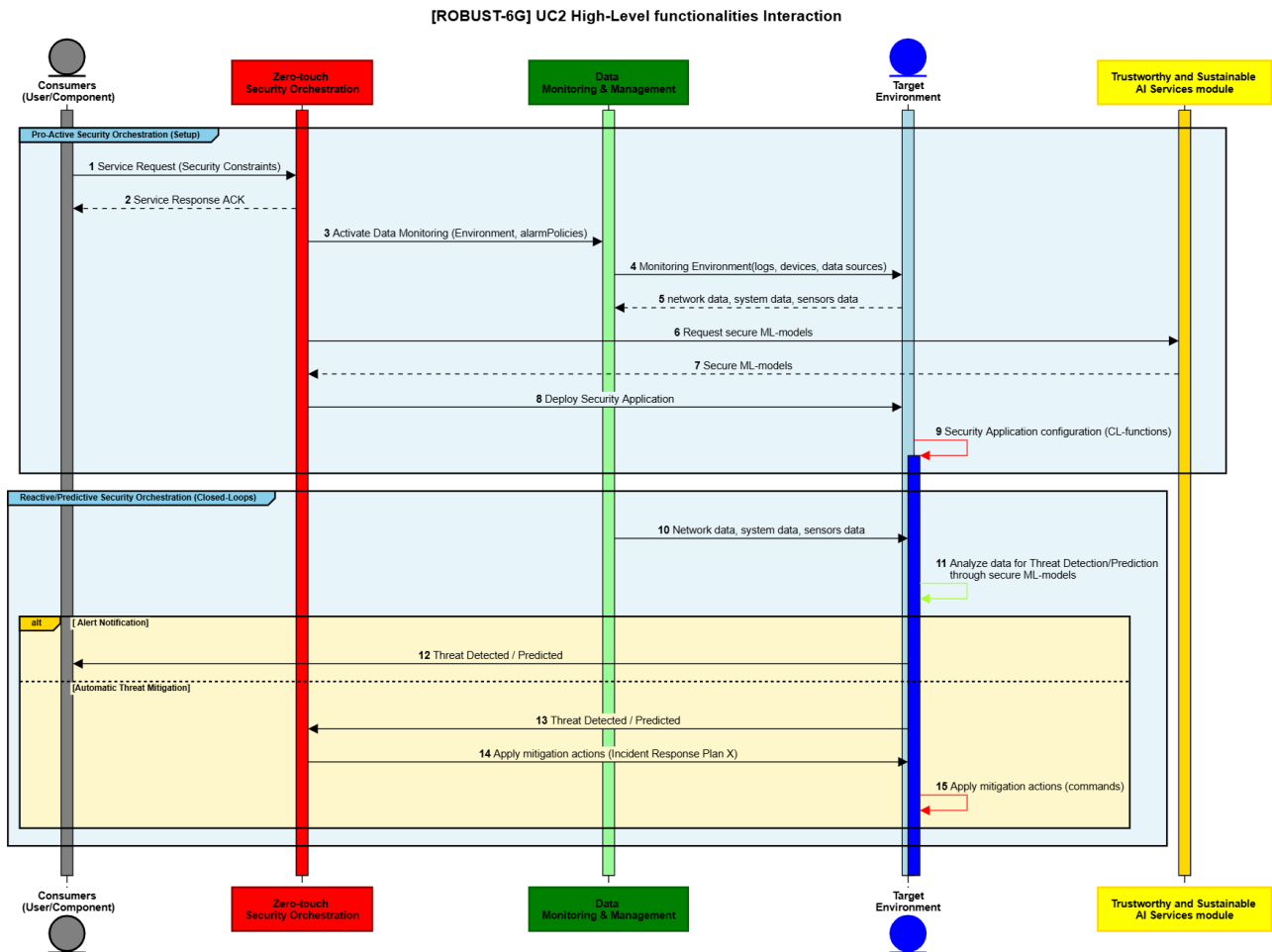


Figure 4-7: UC2 high-level functionalities interactions

The first option is to involve the human by triggering an alert. In sensitive applications, the need for the human-in-the-loop is mandatory due to critical and ethical aspects. In this case, it is the human's responsibility to define additional actions to mitigate the anomaly. On the other hand, for soft applications, it is possible to not involve the human and automatically propose and apply a mitigation plan. This is particularly convenient in applications where, in case of failure, there is minimal damage or significant risks, allowing faster and efficient threat mitigation.

Use Case 3 (UC3) is detailed extensively in Section 2.3. Below, a global sequence diagram illustrates the different interactions between the various components involved. For the sake of readability, and due to the size limitations, this complete diagram is shown in Figure 4-8, Figure 4-9, and Figure 4-10.

The diagram of Figure 4-8 illustrates the authentication and authorization flow within the system. First, the Application initiates an authentication request to the Identity Provider (IdP) containing its identity name and credentials. This corresponding IdP verifies these credentials internally and, upon successful verification, sends the authenticated identity information to the Policy Enforcement Point (PEP). The PEP then initiates an authorization request to the Policy Decision Point (PDP), which evaluates the access rights for the provided identity.

After processing the authorization, the PDP sends the decision back to the PEP, the PEP returns this authorization decision to the IdP, which then issues an access token to the Application.

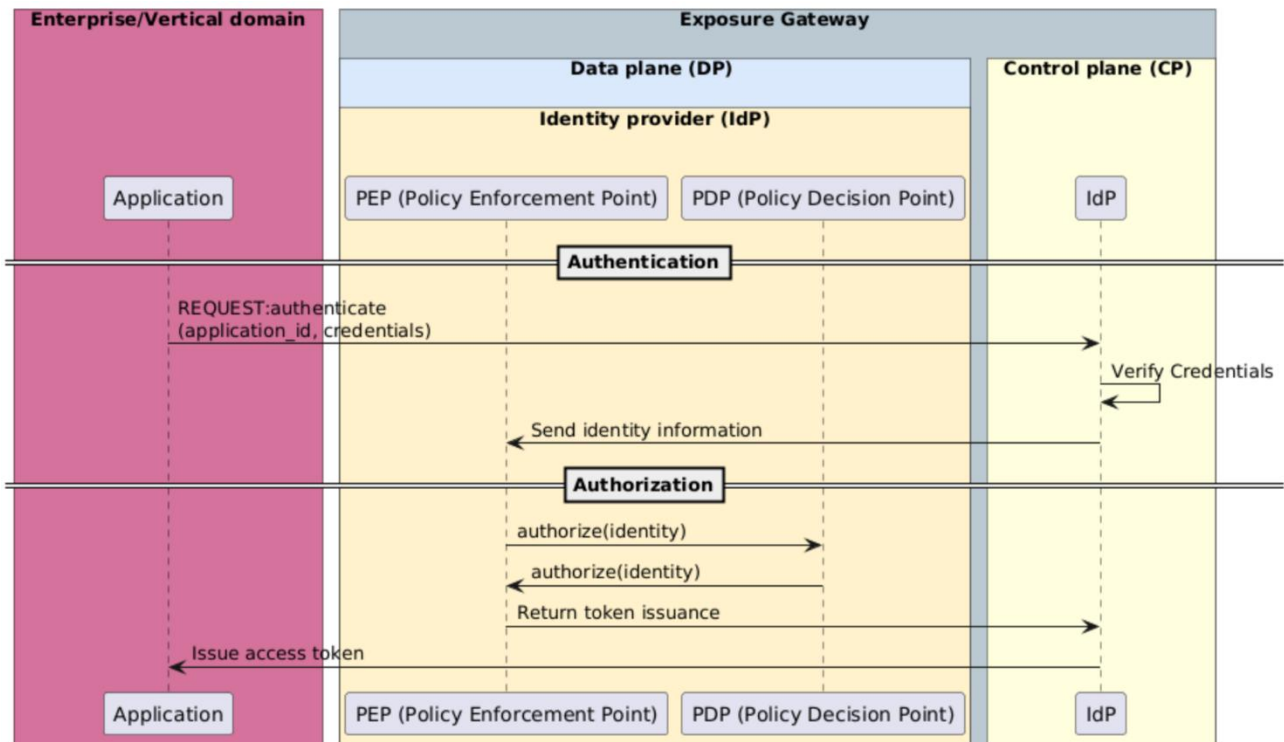


Figure 4-8: UC3-1 interactions

The diagram illustrates the API GET request flow within a system where the Application interacts with various components to retrieve information about a specific service. The flow begins with the Application making a GET request via the API Gateway, which processes this request by validating the access token through the PDP. Once validated, the API Gateway forwards the request to the Transformation Function. Simultaneously, the Programmable Monitoring Platform, the Infrastructure, Network, and Service layer sources continuously send monitoring data to the Data Fabric, where the Knowledge Graph organizes and stores this information. When the Transformation Function receives the GET request, it queries the Data Fabric for relevant monitoring information. The Data Fabric responds with the requested data, which the Transformation Function then returns to the API Gateway. Finally, the gateway relays the data back to the Application, completing the request flow.

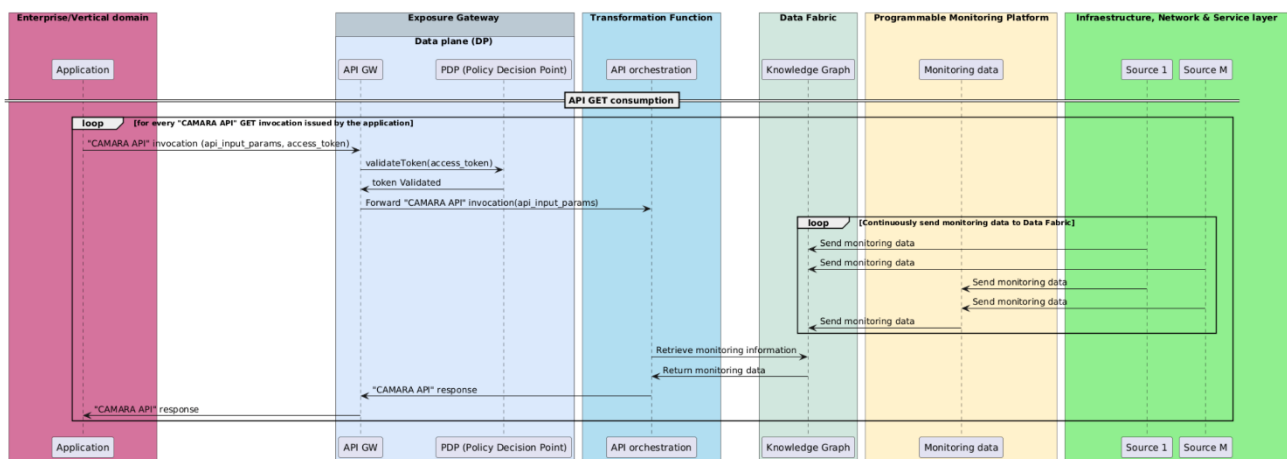


Figure 4-9: UC3-2 interactions

The diagram illustrates the API POST request flow within a system where the Application sends configuration requests. The process begins with the Application issuing a POST request through the API Gateway, including API parameters and an access token. The gateway first validates this token through the PDP. Upon successful validation, the gateway forwards the request to the Transformation Function, which handles configuration actions across various components. The Transformation Function sends configuration instructions as Internal API requests to each Infrastructure, Network, and Service layer source. Each source executes the configuration action and responds back to the Transformation Function. Once all responses are received, the Transformation

Function sends a consolidated response back to the API Gateway, which relays it to the Application, completing the POST request flow.

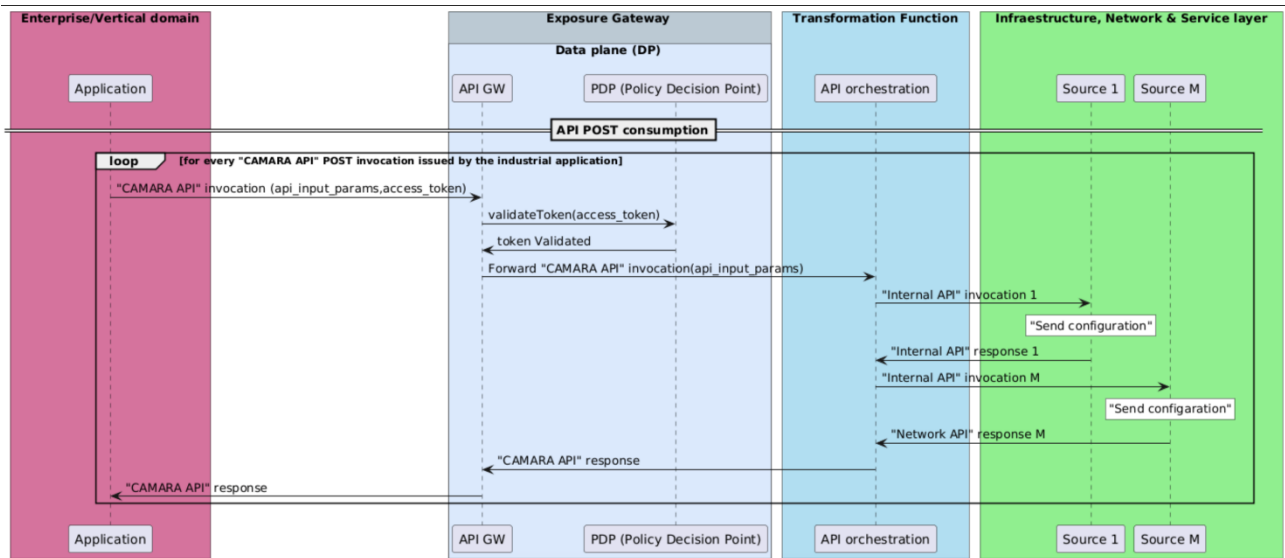


Figure 4-10: UC3-3 interactions

5 ROBUST-6G dataspace

The ROBUST-6G dataspace is built around two core modules: Data Fabric and Data Governance, which together create a secure, efficient, and standardized environment for data management across distributed domains. Designed for the complex and rapidly evolving landscape of 6G networks, this architecture emphasizes interoperability and robust data governance. Figure 5-1 illustrates the dataspace architecture, highlighting the interaction between these modules to form a cohesive, E2E data management platform.

In Figure 5-1, a data consumer accesses what is referred to as a data product. A data product in ROBUST-6G includes data, metadata, and the software necessary for processing, all developed in alignment with Findability, Accessibility, Interoperability, and Reusability (FAIR) principles to ensure the data's optimal usability across multiple contexts. These principles define the characteristics of a ROBUST-6G data product as:

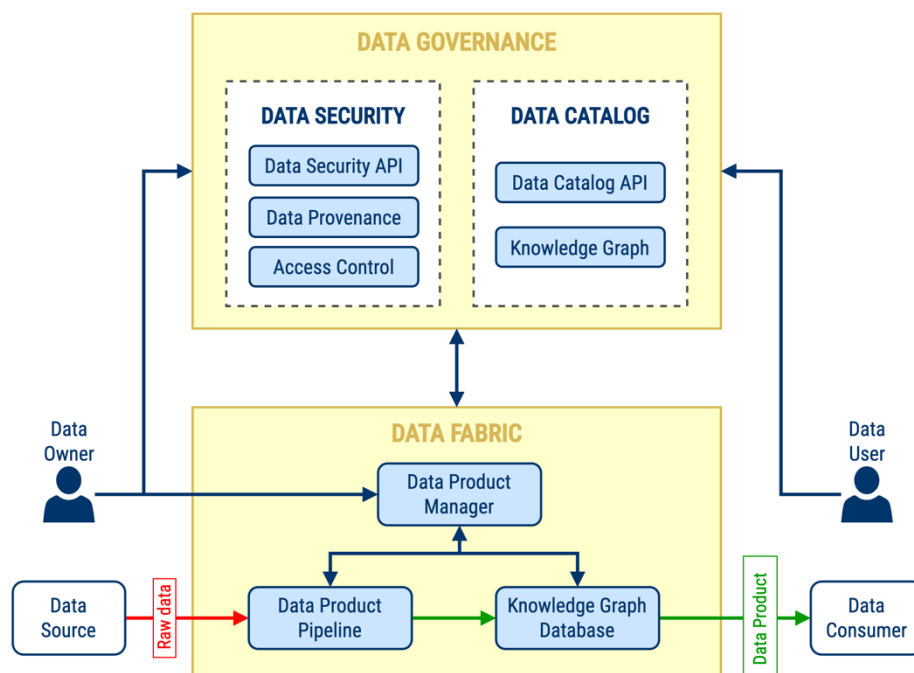


Figure 5-1: Dataspace architecture

1. **Findable:** Data assets are thoroughly tracked, with comprehensive details on their location and ownership, ensuring accountability and discoverability. All data sources are clearly identified for ease of access.
2. **Accessible:** A shared infrastructure enables consistent data access throughout the continuum, with controlled exposure to only authorized consumers, enhancing security and uniform access.
3. **Interoperable:** Agreed-upon data models are leveraged to ensure data can be easily understood by any consumer within the continuum, fostering clarity and consistent data interpretation.
4. **Reusable:** Data is designed for open use and interoperability, allowing seamless application across various UCs and domains, maximizing its value and applicability.

5.1 Data fabric

A cornerstone of the ROBUST-6G architecture is its integration of a knowledge graph within the data fabric. This knowledge graph supports the flexible and dynamic nature of 6G environments, providing a data model that naturally connects diverse data sources through relationships and semantic annotations. The knowledge graph thus enables interoperability and simplifies integration across heterogeneous data types, while allowing for more sophisticated querying and data analysis.

This involves transforming incoming unstructured data into a structured format that organizes it as a collection of well-defined concepts. At this stage, ontologies play a pivotal role, acting as the backbone for establishing relationships between diverse entities and concepts, as shown in Figure 5-2. Ontologies enable raw data to be converted into a structured, semantically rich format that enhances interpretability. This structured representation allows for deeper comprehension, enabling devices to process and recognize real-world concepts and relationships in a meaningful way.

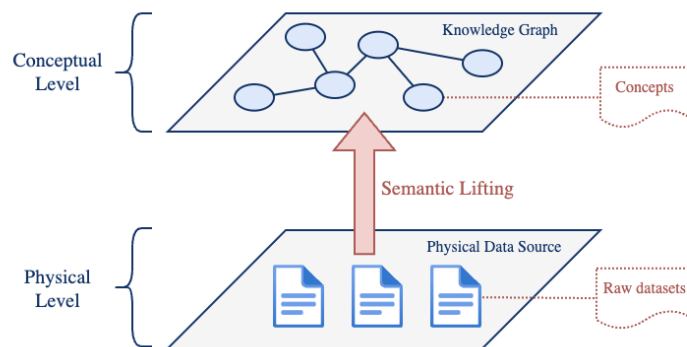


Figure 5-2: Data modelling

The ROBUST-6G Data Fabric is composed of several interconnected components, as illustrated in Figure 5-3:

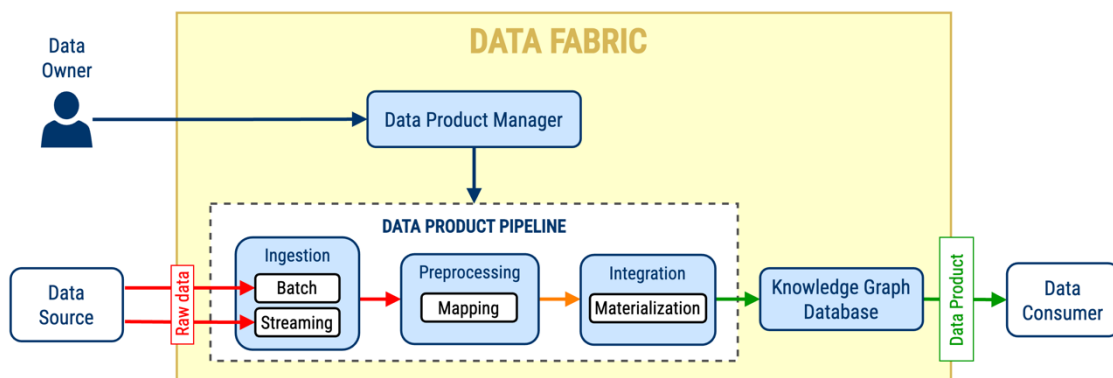


Figure 5-3: Data Fabric

- **Data Product Manager:** Serving as the main interface of the Data Fabric, this component facilitates the onboarding and registration of data products by data product owners. It orchestrates the Data

Product Pipeline for converting raw datasets into data products and collaborates with the Data Catalogue and Data Security components to govern these new data products.

- **Data Product Pipeline:** This pipeline transforms raw datasets into interoperable, semantic data products that are integrated into the knowledge graph. This transformation is unnecessary for “native” data sources that already adhere to a standard format and are semantically annotated according to the agreed-upon ontologies within the continuum. The Data Product Pipeline includes several key subcomponents designed to efficiently process and manage data:
 - **Ingestion:** These modules are responsible for ingesting data from sources listed in the Data Catalog. Collectors can operate in both batch and real-time modes, adapting to the nature of the target data source. For batch data sources, collectors are scheduled to periodically extract data.
 - **Preprocessing:** Transforms raw data from various sources into a unified format, ensuring consistency across datasets. In this step, the raw data is mapped transformed into Resource Description Framework (RDF) according to a target ontology. To this end, the RDF Mapping Language (RML), which is being standardized under the World Wide Web Consortium (W3C) Knowledge Graph Construction (KGC) Community Group [KGC24], will be used.
 - **Serving:** Pre-processed data and metadata are incorporated into a knowledge graph.
- **Knowledge Graph Database:** This core component maintains the Knowledge Graph of the Data Fabric. In scenarios where multiple Data Fabric instances are interconnected (e.g., across multiple ROBUST-6G domains), each database provides a fragment of the global Knowledge Graph of the dataspace.

5.2 Data governance

Data governance within the ROBUST-6G architecture ensures that data is managed with security, compliance, and transparency as top priorities. Key governance components include:

5.2.1 Data Catalog

The Data Catalog building block provides a registry of the data products available within the ROBUST-6G platform. This block provides the pieces for data owners and data governance teams to register data products and oversee their usage. Similarly, the Data Catalog helps data users discover data products based on additional metadata such as terms defined in a glossary business or the people that own these data products. Based on these requirements, the Data Catalog will build upon a metamodel that is drafted in Figure 5-4.

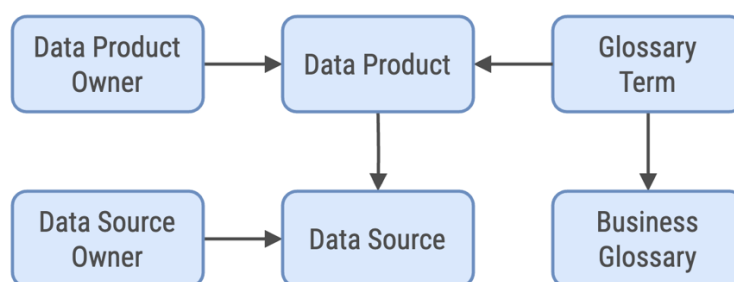


Figure 5-4: Conceptual metamodel of the Data Catalog

In essence, data products are built from data sources by using the Data Fabric. Each data product is published and curated by a data product owner. Similarly, each data source will have a person that accounts for the correct functioning and maintenance of it. The Data Catalog will include business glossaries containing terms that have been formally defined among members of the ROBUST-6G platform. In turn, these terms will be linked to data products based on the concepts that the contained data refers to.

To support this, the Data Catalog block is built upon two main components: the Data Catalog API and the Data Catalog Knowledge Graph.

Data Catalog API

This component implements an API that introduces an abstraction layer for data owners and data users to interact with the Data Catalog. At a high level, this API exposes services for managing the following core concepts in the Data Catalog metamodel:

- **Data Source:**
 - Registration data sources, containing raw data, from which data products are built. Examples of data sources are Relational Database Management Systems (RDBMS), message queue systems (Kafka, Message Queuing Telemetry Transport–MQTT) or remote files and APIs.
 - Registration knowledge graph databases of the different data fabric, to be managed as data sources of data products.
- **Data Product:**
 - Registration of data products that have been built using the Data Fabric.
 - Search for data products and provide information about the Data Fabric where they can be accessed (i.e., the knowledge graph database of the Data Fabric).
- **Business Glossary:**
 - Creation of a business glossary.
 - Search of terms defined in the glossary.
 - Assignment of glossary terms to data products.
- **Ownership:**
 - Assignment of owners to data products.

Data Catalog Knowledge Graph

The Data Catalog internally stores all metadata in a knowledge graph. Deriving from the metamodel depicted in Figure 5-5, the Data Catalog structures the metadata based on a custom ontology named Data Catalog Ontology. The development of this new ontology is currently conducted following the guidelines defined by the Linked Open Terms (LOT) methodology [PFF+22]. In this regard, the reuse of existing standard ontologies that can be leveraged is still under exploration.

The standard Data Catalog Vocabulary (DCAT) 3.0 Ontology [DCAT24], which is broadly used by official institutions like the European Commission, could serve as the basis for the Data Catalog ontology as it defines core concepts like catalog, dataset, and data service. Combined with DCAT Ontology, the SPARQL Protocol and RDF Query Language (SPARQL) 1.2 Service Description Ontology (SPARQL-SD) [SPARQL24] can be leveraged to represent the SPARQL endpoint for the knowledge graph of the Data Fabric, where the data products are available. A good starting point for combining both DCAT and the SPARQL-SD is suggested [Ate16], where `dcat:Dataset` and `sd:Dataset` concepts are merged so as to manage graphs or named graphs as datasets in the Data Catalog (see Figure 5-5). In this sense, a data product created in the Data Fabric could be represented within a named graph, which, in turn, is registered as a dataset in the Data Catalog.

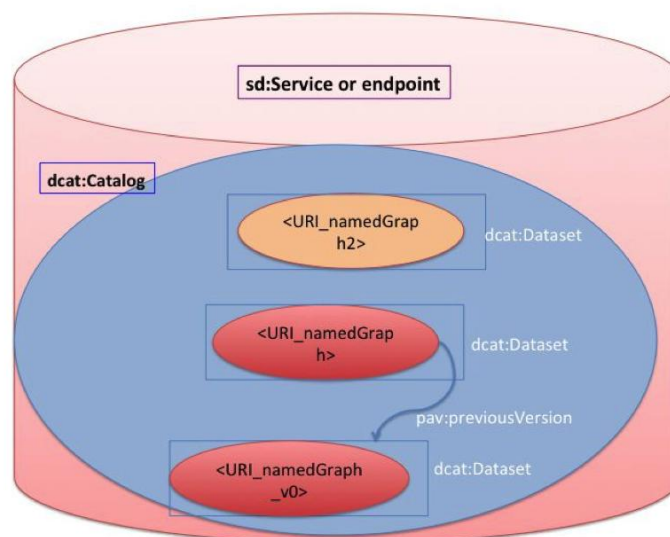


Figure 5-5: Combination of DCAT and RDF datasets (source: [Ate16])

On the other hand, the Data Catalog must cope with the **business glossary terms** that link to the data products registered in the catalog. In this sense, the Simple Knowledge Organization System Ontology (SKOS) [MB08] is the go-to standard ontology for knowledge management. This ontology allows for building glossaries, taxonomies, and thesauri. The details on the combination of business terms defined in a glossary with datasets from a data catalog are already suggested by the DCAT specification.

The data products registered in the data catalog are the responsibility of **data owners**, therefore, the Data Catalog must capture information about the members and teams of an organization, and their ownership of data products. Here the Organization Ontology (ORG) [ORG14] and the Friend-of-a-Friend Ontology (FOAF) [BM14] have been identified as the best candidates. As in the case of the business glossary, the DCAT ontology specification also provides guidelines for capturing the responsibilities of organization members with respect to the data products of a data catalog.

Lastly, the Data Catalog should ideally support data quality to provide trust in the data products. To this end, the data catalog can integrate **data lineage** in the data catalog so to keep track of how data products have been generated in the data fabric. In this sense, the RML ontology itself, which is used in the Data Product Pipeline, can contribute to the lineage of the data. Similarly, the standard PROV Ontology (PROV-O) [PROVO13] could be combined with RML to provide further lineage information, though this possibility remains to be discussed with the W3C KGC, which has defined the RML ontology.

Based on this analysis, the following releases of the Data Catalog block will bring a consolidated Data Catalog Ontology as well as the mechanisms that enable the population of these metadata in the knowledge graph of the Data Catalog. For the implementation of such mechanisms, the use of the RML language is also foreseen, as done for the creation of data products in the Data Fabric.

5.2.2 Data Security

In the Data Security module, we have three key functionalities:

- **Access Control:** Ensures that data is only accessible to authorized users or systems by implementing fine-grained permissions and authentication mechanisms.
- **Data Provenance:** Tracks the origin, history, and lifecycle of data, providing transparency and accountability for all actions performed on it.
- **Data Security API:** Offers an interface to manage and enforce security policies, enabling consistent policy application and updates across the system.

These functionalities work in synergy to provide robust, transparent, and efficient protection for your data.

Access Control

The Open Authorization (OAuth) framework serves as the reference standard for access control mechanisms, recognized for its effectiveness and widespread adoption. This study aligns with the objectives of the IETF Workload Identity in Multi-Service Environments (WIMSE) working group, which addresses the complexities of implementing fine-grained, least-privilege access control for workloads deployed across multiple service platforms.

To establish the ROBUST-6G access control framework, several essential components must be considered:

- **Resources:** These include all protected assets that users may wish to access, encompassing physical items, virtual entities, and elements related to services and applications.
- **Users:** These are individuals or entities that interact with the system, either by requesting access to resources or by modifying policies to align with operational requirements. Notably, a network entity may function as either a resource or a user depending on the context of specific access decisions.
- **Identity Provider (IdP):** This centralized service manages the identities of all participating entities, typically utilizing protocols such as Lightweight Directory Access Protocol (LDAP). The IdP ensures consistent identity verification across multiple systems and services.
- **Policy Enforcement Point (PEP):** The PEP is responsible for intercepting access requests and enforcing trust and access policies. It guarantees that only authorized users can gain access to the requested resources. The specific policies that the PEP enforces are defined by the PDP.
- **Policy Decision Point (PDP):** The PDP evaluates incoming access requests against established policies to determine whether access should be granted or denied. It sets the rules and criteria for access control based on user roles, attributes, and contextual factors.

- **Policy Administration Point (PAP):** This component is tasked with creating, managing, and modifying access policies. The PAP facilitates dynamic adjustments based on the evolving needs of data owners or environmental conditions.
- **Accounting Ledger (AL):** The AL meticulously tracks all access control activities, ensuring that every action is logged for auditing and compliance purposes. It maintains a comprehensive record of decisions made over time, allowing for thorough review and analysis.

The following architecture diagram illustrates in the components and their interactions within the access control framework.

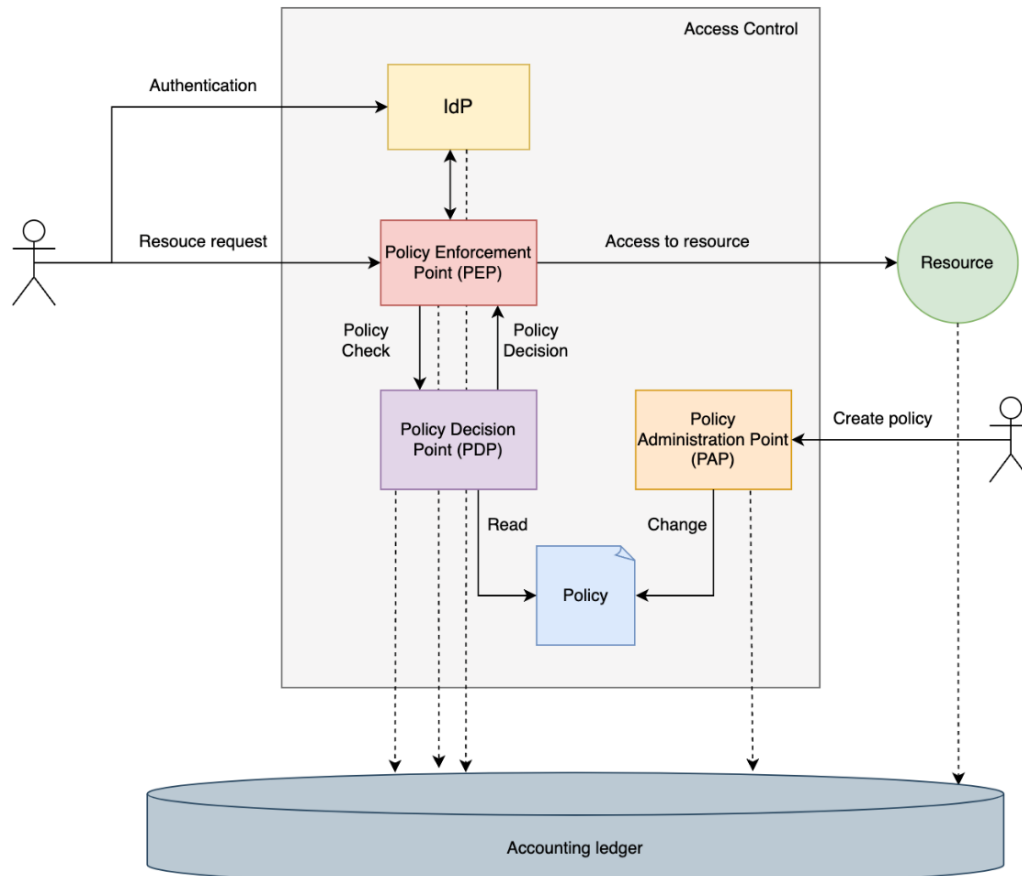


Figure 5-6: Access control mechanism

Data flow management ensures that access requests are handled efficiently and securely from initial authentication through to final authorization and activity logging, as illustrated in Figure 5-7.

The access request process begins when a user submits their credentials to the IdP for verification. Upon successful authentication, the IdP transmits the user's information to the PEP, which subsequently queries the PDP for the applicable access policies. The PDP evaluates these policies and returns an access decision to the PEP.

If access is granted, the PEP instructs the IdP to issue an access token to the user. The user then presents this token to the PEP when requesting access to a specific resource. The PEP forwards the access request to the PDP, which assesses it against the defined policies. The PDP's decision regarding the request is then communicated back to the PEP.

Should access be granted, the PEP enables the user to interact with the resource. Conversely, if access is denied, the PEP promptly informs the user. This structured flow guarantees that access decisions are made based on thorough policy evaluations and user authentication, ensuring both security and clarity in the access control process.

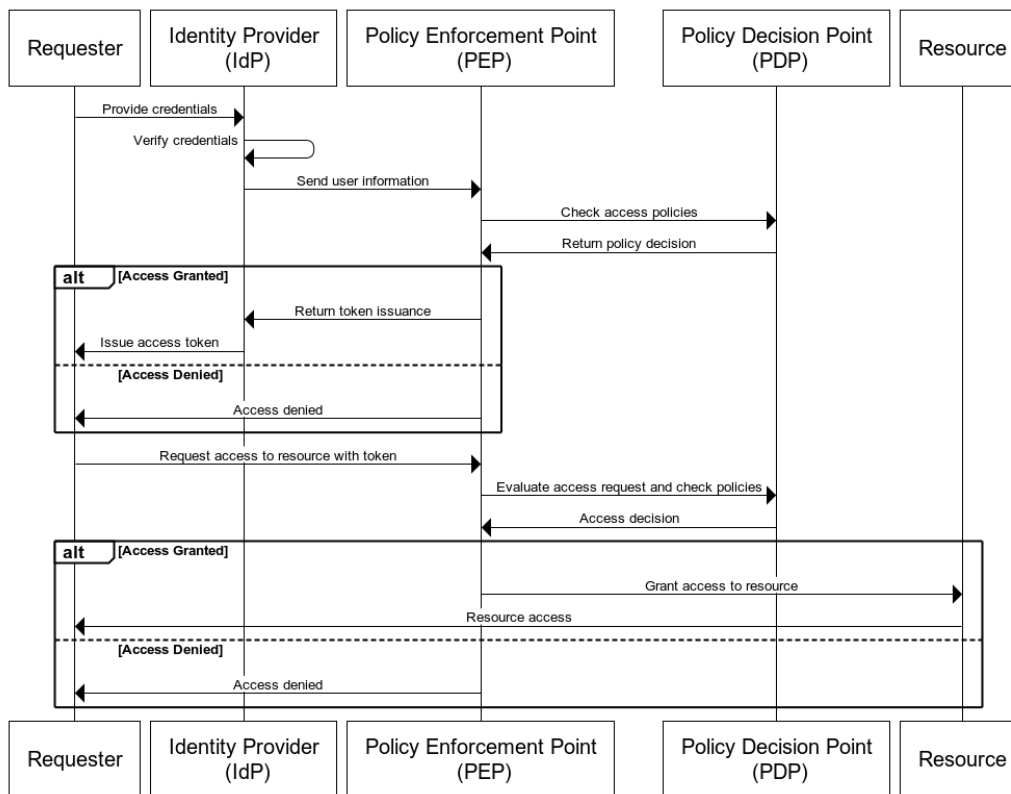


Figure 5-7: Data flow for access control mechanism

Data Security API

Access control is enforced through the Data Security API, which operates under the ROBUST-6G's Authentication, Authorization, and Accounting (AAA) framework. This component ensures that only authenticated and authorized users can access data, supporting granular policy enforcement. It strengthens compliance with privacy, regulatory, and security standards by enabling data product owners and governance teams to define precise, context-specific access controls.

Data product owners and members of the data governance team can define granular access control policies based on various patterns, such as roles or attributes. The Data Security service handles the registration, management, and enforcement of these policies, ensuring that data access remains secure and governed appropriately.

The concept of Policy-as-Code (PaC) is essential in modernizing access control management. Using declarative languages like Rego [Rego24], authorization policies can be managed programmatically across multiple systems and environments. PaC offers significant advantages over traditional manual methods by enabling agile, scalable, and secure policy management. This approach not only simplifies policy updates and enforcement but also enhances the overall security posture of the system. Rego [Rego24] is a policy language specifically designed to define rules over complex data structures. Drawing inspiration from Datalog [Wil24], Rego offers a powerful set of features that make it ideal for modern, data-driven applications:

- **Declarative Syntax:** Policies are defined in a clear, human-readable format.
- **Expressiveness:** Supports nested rules, structured data, and built-in functions for granular control over user actions and resources.
- **Interoperability:** Operates seamlessly on JavaScript Object Notation (JSON) data, making it highly effective for distributed and cloud-native applications.

As a reference, an example Rego policy is included below:

```

package example
# Allow read access to resources if the user has the "read" role
default allow = false
allow {
    input.user.role == "read" }
    
```

In this example, the policy defines access rules based on user roles. The rule evaluates input data, returning a Boolean value that dictates whether the action is permitted. Rego allows this logic to be extended to extremely detailed conditions, enabling fine-grained control over which users can perform specific actions, on which resources, and under which circumstances.

The Data Security API provides a RESTful interface to manage and enforce Rego-based policies efficiently.

Key Features

- **Policy Registration:** Easily upload new access control policies using .rego files.
- **Retrieve Policies:** Fetch the content of a specific policy, with the option to download it as a .rego file.
- **List Policies:** View all registered policies in a single request.
- **Update Policies:** Modify existing policies without re-registration.
- **Delete Policies:** Securely remove policies via API calls.
- **Granular Control Enforcement:** Ensure policies are precise and enforceable at scale.

Figure 5-8 shows the FastAPI interface, which provides a user-friendly platform to interact with the Data Security API:

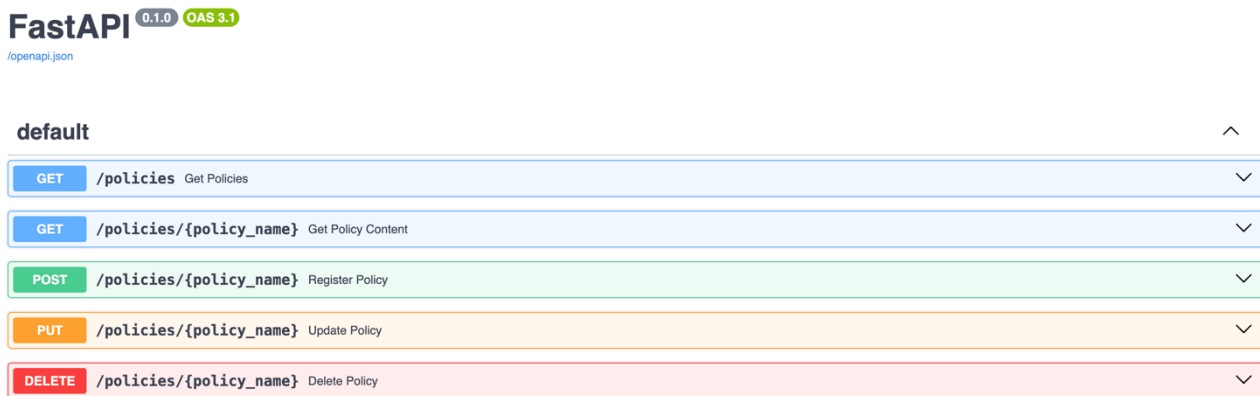


Figure 5-8: Data Security API

Data Provenance

To establish provenance, ensuring both the origin and integrity of data, we propose implementing digital signatures in line with the IETF draft on Concise Binary Object Representation (CBOR) Object Signing and Encryption (COSE) [LPF+23]. This approach provides a robust method for validating the authenticity and integrity of data by embedding digital signatures within it. Leveraging COSE, signatures are generated and verified according to Public Key Infrastructure (PKI) principles.

The process consists of several key steps [LPF+23]:

1. **Canonicalization:** Before signature generation, the content undergoes canonicalization, guaranteeing a consistent data representation regardless of serialization format.
2. **Signature Generation:** Using the COSE Sign1 structure, the signature is generated with components including the key ID (kid), serialization method, algorithm parameters, and the signature itself.
3. **Signature Verification:** The generated signature is then validated against “externally supplied data”, specifically the content used during signature creation.

It is crucial to note that digital signatures alone do not prevent data tampering if the data is compromised before signing. Thus, this solution is complemented by access control mechanisms to safeguard against unauthorized access and further reinforce data integrity.

Figure 5-9 illustrates the process used to ensure provenance through digital signatures. By combining these robust governance and security frameworks, the ROBUST-6G dataspace architecture provides a resilient and adaptable environment for managing, securing, and accessing data in the evolving landscape of 6G. This approach guarantees that data products within the continuum are both trusted and versatile, empowering a broad spectrum of secure data UCs.

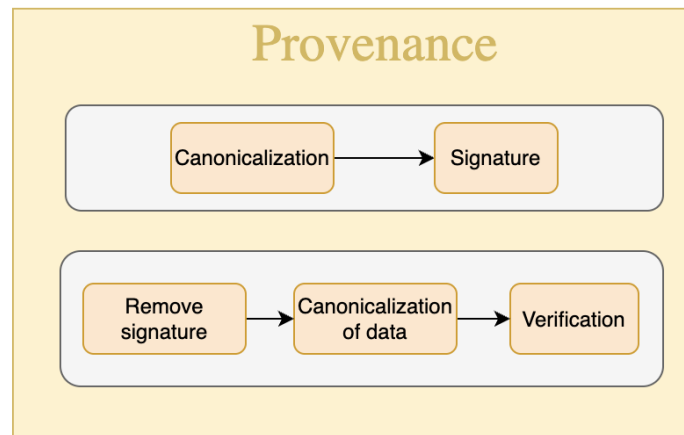


Figure 5-9: Provenance

6 Conclusion

The ROBUST-6G project aims to provide an integrated approach of smart security services for 6G networks through the exploitation of distributed and trusted AI/ML, zero-touch integrated security management and orchestration mechanisms, AI/ML-driven physical layer security technologies, pervasive monitoring, and effective data management capabilities. In this context, this deliverable has presented the functional and technical requirements, applicability UCs targeted by the project, initial version of the system architecture as well as the ROBUST-6G dataspace.

Three UCs are considered by the project. The first one deals with the trustworthiness assessment of AI/ML models in distributed 6G networks using decentralized federated learning. To strengthen the trustworthiness of the AI/ML models, different aspects such as robustness, sustainability, explainability, fairness, privacy and security in both physical and sensing layers are analysed. The UC also includes interactions between nodes of different hierarchies (cloud, edge, and extreme edge) to generate shared models while preserving privacy, assessing reputation and mitigating threats that could impact the AI/ML models. Two scenarios have been highlighted to build, on the one hand, a DFL agnostic framework for trusted AI/ML models and, on the other hand, to obtain and evaluate physical and security measures.

The second UC aims at demonstrating the capabilities of ROBUST-6G security orchestration and automation solutions with focus on anomalies and attacks into smart IoT environments. The UC includes three scenarios considering different attacks on IoT devices and platform sited at far/extreme edge along with intended detection and remediation plans based on AI-driven closed-loops.

The third one aims at extending the Open Gateway framework with the advanced security capabilities of ROBUST-6G, introducing the concept of Network-Security-as-a-Service (NetSecaaS). This UC focuses on enabling application developers and enterprises to seamlessly apply security capabilities through novel APIs developed within the CAMARA project, which abstract the complexity of telecommunications and facilitate security management.

This deliverable has also presented the initial versions of the high-level and functional architectures of the ROBUST-6G project together with the deployment view of the developed solutions. These are built with a set of capabilities in mind, namely: distributed and trustworthy AI, exposure of security services, zero-touch orchestration and automation for security services, intelligent physical layer security solutions and effective data management & governance capabilities.

The integration of the abovementioned services has led to the design of the architecture around the ROBUST-6G dataspace, built upon the foundational modules of Data Fabric and Data Governance. These modules work in tandem to create a secure, efficient, and standardized environment for data management, tailored to the needs of distributed domains within the 6G ecosystem. The architecture emphasizes interoperability, enabling seamless integration across diverse systems, and robust data governance to ensure data quality, privacy, and secure access.

References

- [Ate16] G. A. Atemez, “Applying DCAT vocabulary on RDF datasets”, November 2016, https://www.w3.org/2016/11/sdsvoc/SDSVoc16_paper_6.
- [BKT+22] A. Borys, A. Kamruzzaman, H. N. Thakur, J. C. Brickley, M. L. Ali and K. Thakur, “An evaluation of IoT DDoS cryptojacking malware and Mirai botnet”. In Proceedings of the 2022 IEEE World AI IoT Congress, pp. 725-729, June 2022.
- [BM14] D. Brickley and L. Miller, “FOAF vocabulary specification 0.99”, Namespace Document, January 2014, <http://xmlns.com/foaf/spec>.
- [BM21] N. Bouacida and P. Mohapatra, “Vulnerabilities in federated learning”. IEEE Access, vol. 9, pp. 63229-63249, April 2021.
- [Bra97] S. Bradner, “RFC 2119: Key words for use in RFCs to indicate requirement levels”, March 1997, <https://www.ietf.org/rfc/rfc2119.txt>
- [Cam23] The CAMARA Project, “The telco global API alliance”, 2023, <https://camaraproject.org>.
- [CCM21] E. Catak, F. O. Catak and A. Moldsvor, “Adversarial machine learning security problems for 6G: mmWave beam prediction use-case”. In Proceedings of the 2021 IEEE International Black Sea Conference on Communications and Networking, pp. 1-6, May 2021.
- [Chr24] The 6G-CHRONOS Project, “AI-assisted beyond 5G-6G architecture with deterministic networking for industrial communications”, 2022-2024, <https://wimUNET.UGR.es/projects/6gchronos.php>.
- [CVH24] J. Christ, L. Visengeriyeva and S. Harrer, “Data mesh architecture: Data mesh from an engineering perspective”, 2024, <https://www.datamesh-architecture.com>.
- [CZD23] H. Chi, Q. Zeng and X. Du, “Detecting and handling IoT interaction threats in multi-platform multi-control-channel smart homes”. In Proceedings of the 32nd USENIX Security Symposium (USENIX Security 23), pp. 1559-1576, August 2023.
- [Dat24] The 6G-DATADRIVEN Project, “Data driven sustainable next generation (B5G and 6G) networks for manufacturing and emergency response”, 2023-2024, <https://unica6g.it.uc3m.es/en/6g-datadriven>.
- [DCAT24] World Wide Web Consortium, “Data Catalog Vocabulary (DCAT) - Version 3”, W3C Recommendation, August 2024, <https://www.w3.org/TR/vocab-dcat-3>.
- [Deh22] Z. Dehghani, “Data mesh: Delivering data-driven value at scale”. O’Reilly Media, April 2022.
- [Gar24] Gartner Inc., “Using data fabric architecture to modernize data integration”, 2024, <https://www.gartner.com/en/data-analytics/topics/data-fabric>.
- [Gsm23] GSMA, “The ecosystem for Open Gateway NaaS API development”, White Paper, June 2023, https://www.gsma.com/solutions-and-impact/gsma-open-gateway/gsma_resources/naas-ecosystem-whitepaper.
- [GSS23] V. Gugueoth, S. Safavat and S. Shetty, “Security of Internet of Things (IoT) using federated learning and deep learning—Recent advancements, issues and prospects”. ICT Express, vol. 9, no. 5, pp. 941-960, October 2023.
- [GYZ+21] F. Guo, F. R. Yu, H. Zhang, X. Li, H. Ji and V. C. Leung, “Enabling massive IoT toward 6G: A comprehensive survey”. IEEE Internet of Things Journal, vol. 8, no. 15, pp. 11891-11915, August 2021.
- [Hex24-D33] Hexa-X-II project consortium, “Deliverable D3.3: Initial analysis of architectural enablers and framework”, April 2024.
- [HLB+18] R. Heartfield, G. Loukas, S. Budimir, *et. al.*, “A taxonomy of cyber-physical threats and impact in the smart home”. Computers & Security, vol. 78, pp. 398-428, September 2018.

- [JOR+23] O. Jullian, B. Otero, E. Rodriguez, N. Gutierrez, H. Antona and R. Canal, “Deep-learning based detection for cyber-attacks in IoT networks: A distributed attack detection framework”. *Journal of Network and Systems Management*, vol. 31, art. no. 33, pp. 1-24, February 2023.
- [JSG+22] J. M. Jorquera Valero, P. M. Sánchez Sánchez, M. Gil Pérez, A. Huertas Celdrán and G. Martínez Pérez, “Toward pre-standardization of reputation-based trust models beyond 5G”. *Computer Standards & Interfaces*, vol. 81, art. no. 103596, pp. 1-17, April 2022.
- [K3s24] K3s Project Authors, “Lightweight Kubernetes”, 2024, <https://k3s.io>.
- [K8s24] The Kubernetes Authors, “Kubernetes: an open-source system for automating deployment, scaling, and management of containerized applications”, 2024, <https://kubernetes.io>.
- [KAK+23] N. W. Khan, M. S. Alshehri, M. A. Khan, *et al.*, “A hybrid deep learning-based intrusion detection system for IoT networks”. *Mathematical Biosciences and Engineering*, vol. 20, no. 8, pp. 13491-13520, June 2023.
- [KCC+23] M. Kuzlu, F. O. Catak, U. Cali, E. Catak, and O. Guler, “Adversarial security mitigations of mmWave beamforming prediction models using defensive distillation and adversarial retraining”. *International Journal of Information Security*, vol. 22, no. 2, pp. 319-332, April 2023.
- [KGC24] W3C Community Development Team, “Knowledge graph construction community group”, 2024, <https://www.w3.org/community/kg-construct>.
- [KMA+21] P. Kairouz, H. B. McMahan, B. Avent, *et al.*, “Advances and open problems in federated learning”. *Foundations and Trends in Machine Learning*, vol. 14, no. 1-2, pp. 1-210, June 2021.
- [LPF+23] D. Lopez, A. Pastor, A. H. Feng, H. Birkholz and S. Garcia, “Applying COSE signatures for YANG data provenance”. IETF Internet-Draft draft-lopez-opsawg-yang-provenance-03, July 2024, <https://datatracker.ietf.org/doc/draft-lopez-opsawg-yang-provenance/03>.
- [MB08] A. Miles and S. Bechhofer, “SKOS Simple knowledge organization system RDF schema”, August 2008, <https://www.w3.org/TR/2008/WD-skos-reference-20080829/skos.html>.
- [MHE+16] D. Meyer, J. Haase, M. Eckert and B. Klauer, “A threat-model for building and home automation”. In *Proceedings of the 2016 IEEE 14th International Conference on Industrial Informatics*, pp. 860-866, July 2016.
- [MJC+21] L. Mucchi, S. Jayousi, S. Caputo, *et al.* “Physical-layer security in 6G networks”. *IEEE Open Journal of the Communications Society*, vol. 2, pp. 1901-1914, August 2021.
- [MKP+22] V. Mothukuri, P. Khare, R. M. Parizi, S. Pouriye, A. Dehghantanha and G. Srivastava, “Federated-learning-based anomaly detection for IoT security attacks”. *IEEE Internet of Things Journal*, vol. 9, no. 4, pp. 2545-2554, February 2022.
- [OD22] J. Ordonez-Lucena and F. Dsouza, “Pathways towards network-as-a-service: the CAMARA project”. In *Proceedings of the ACM SIGCOMM Workshop on Network-Application Integration*, pp. 53-59, August 2022.
- [ORG14] World Wide Web Consortium, “The organization ontology”, W3C Recommendation, January 2014, <https://www.w3.org/TR/vocab-org>.
- [OS24] OpenInfra Foundation, “OpenStack: an open-source cloud computing standard to support virtual machines, container and bare metal workloads”, 2024, <https://www.openstack.org>.
- [PFF+22] M. Poveda-Villalón, A. Fernández-Izquierdo, M. Fernández-López and R. García-Castro, “LOT: An industrial oriented ontology engineering framework”. *Engineering Applications of Artificial Intelligence*, vol. 111, art. no. 104755, pp. 1-22, May 2022.
- [PROVO13] World Wide Web Consortium, “PROV-O: The PROV ontology”, W3C Recommendation, April 2013, <https://www.w3.org/TR/prov-o>.
- [Rego24] Open Policy Agent, “Policy Language: the native query language Rego”, 2024, <https://www.openpolicyagent.org/docs/latest/policy-language>.
- [Snort24] Cisco, “Snort: an open-source intrusion prevention system”, 2024, <https://www.snort.org>.

- [SPARQL24] World Wide Web Consortium, “SPARQL 1.2 service description”, W3C Working Draft, November 2024, <https://www.w3.org/TR/sparql12-service-description>.
- [TAU22] E. Tekiner, A. Acar and A. S. Uluagac, “A lightweight IoT cryptojacking detection mechanism in heterogeneous smart home networks”. Network and Distributed System Security (NDSS) Symposium, pp. 1-15, April 2022.
- [Ton24] 5TONIC, “An open research and innovation laboratory focusing on 5G technologies”, 2024, <https://www.5tonic.org>.
- [URB+21] M. A. Uusitalo, P. Rugeland, M. R. Boldi, *et al.*, “6G vision, value, use cases and technologies from European 6G flagship project Hexa-X”. IEEE Access, vol. 9, pp. 160004-160020, November 2021.
- [Wil24] M. Willsey (UC Berkeley), “Declarative program analysis and optimization (CS294-260): Datalog”, 2024, <https://inst.eecs.berkeley.edu/~cs294-260/sp24/2024-02-05-datalog>.
- [WL24] W. Wei and L. Liu, “Trustworthy distributed AI systems: Robustness, privacy, and governance”. ACM Computing Surveys, Just Accepted, pp. 1-38, January 2024.
- [ZLQ+20] Y. Zhan, P. Li, Z. Qu, D. Zeng and S. Guo, “A learning-based incentive mechanism for federated learning”. IEEE Internet of Things Journal, vol. 7, no. 7, pp. 6360-6368, January 2020.
- [ZMZ+22] M. Zolotukhin, P. Miraghaie, D. Zhang, T. Hämäläinen, W. Ke and M. Dunderfelt, “Black-box adversarial examples against intelligent beamforming in 5G networks”. In Proceedings of the 2022 IEEE Conference on Standards for Communications and Networking, pp. 64-70, November 2022.
- [ZSM009-1] ETSI GS ZSM 009-1 Version 1.1.1, “Zero-touch network and Service Management (ZSM); Closed-loop automation; Part 1: Enablers”, 2023-01.