6GSNS
Co-funded by the European Union

SAFE-6G

SAFE-6G

A Smart and Adaptive Framework for Enhancing Trust in 6G Networks

Deliverable D2.4: 6G Trustworthiness KPI & KVI definition and validation methodology

Date: 22/12/2025

Version: v1.1

DISCLAIMER

Grant Agreement	101139031
Document number	D2.4
Document title	6G Trustworthiness KPI & KVI definition and validation methodology
Lead Beneficiary	Telefonica Innovación Digital (TID)
Editor(s)	Daniel García Sánchez (TID) Álvaro Andrés Anaya (TID) Rodrigo Sanz Sanz (TID)
Author(s)	Daniel García Sánchez (TID) Rodrigo Sanz Sanz (TID) Álvaro Andrés Anaya (TID) Alejandro Fornés (UPV) Joaquín Cáceres (ATOS) Sonia Castro (ATOS) Charles Bailly (IMM) Panos Karkazis (UNIWA) Dimitris Uzunidis (UNIWA) Stamatia Drampalou (UNIWA) Spyridon Georgoulas (NCSRD) Vasiliki Rentoula (NCSRD) Ilias Alexandropoulos (NCSRD) Harilaos Koumaras (NCSRD) Nikolaos Zompakis (8BELLS) Vagelis Anagnostopoulos (INF) Eugenia Vergi (INF) George Koumaras (INF) Gaëtan Pruvost (THA) Stéphane Lorin (THA) Thodoris Ioannidis (IQBT) Henry Faure-Geors (KEY) Guillaume Hébert (KEY) Apostolis Garos (SHG) Zouzias Dimitris (eBOS)
Dissemination level	Public
Contractual date of delivery	28/02/2025
Status	Final
File name	SAFE-6G_D2.4_v1.1.pdf

Revision History

Version	
v0.1	Initial table of contents proposal.
v0.2	First KPI contributions and KPI and KVI methodology.
v0.3	Included feedback after 3rd Plenary Meeting.
v0.4	Included feedback from Trust Function responsible after the First Review.
v0.5	Included KVI selection based on the comments from the Second Review performed by the Technical Steering Committee.
v1.0	Final review performed by the Project Coordinator and the Editor. Final version following the Quality check.
v1.1	Updated version after mid-term review

GLOSSARY

Abbreviations/Acronym	Description
AFs	Application Functions
AI	Artificial Intelligence
AMF	Access Management Function
API	Application Programming Interface
AR	Augmented Reality
B5G	Beyond 5G
CI/CD	Continuous Integration/Continuous Delivery
cLoT	calibrated Level of Trustworthiness
CNCF	Cloud Native Computing Foundation
DID	Decentralized Identifiers
DLT	Distributed Ledger Infrastructure
DP	Differential Privacy
DPI	Data Protection Index
EC	European Commission
EOSC	European Open Science Cloud
FL	Federated Learning
GDPR	General Data Protection Regulation
HA	High Availability
KPIs	Key Performance Indicators
KVIs	Key Value Indicators
KVs	Key Values
LoTw	Level of Trustworthiness
ML	Machine Learning
MNO	Mobile Network Operator
N/A	Not available/applicable
NDT	Network's Digital Twin
NF	Network Function
nLoT	non-calibrated Level of Trustworthiness
NLP	Natural Language Processing
NLU	Natural Language Understanding
NRF	Network Repository Function
NSAAR	Network Safe Actions Auto Ratio
NSN	Network Service Node
OSC	Open-Source Community
PS	Privacy Score
QoE	Quality of Experience
QoS	Quality of Service
RBO	Regulatory Body Organization
RFS	Resilience Function Score
RPE	Resource Provisioning Efficiency
RURR	Relative Utilization of Resource Reduction
SBA	Service Based Architecture
SDG	Sustainable Development Goals
SDO	Standard Developing Organization
SDP	Software Defined Perimeter

SFC	Service Function Chaining
SSI	Self-Sovereign Identity
SSP	Security Service Provider
TFs	Trust Functions
TMV WG	Technology Monitoring and Validation Working Group
TV	Telecom Vendor
UN	United Nations
UPF	User Plane Function
USN	User Service Node
VC	Verifiable Credential
VR	Virtual Reality
XAI	eXplainable AI
XR	Extended Reality

EXECUTIVE SUMMARY

This document is a key output of the SAFE-6G project, titled "A Smart and Adaptive Framework for Enhancing Trust in 6G Networks." It focuses on the definition and validation methodology for Key Performance Indicators (KPIs) and Key Value Indicators (KVIs), ensuring the effective measurement and assessment of trust, security, privacy, resilience, and reliability in future 6G networks. The work presented here aligns with the objectives outlined in the project's Grant Agreement (101139031) and contributes to the European Commission's vision of secure, efficient, and user-centric next-generation communication infrastructures.

The document provides a structured approach to defining KPIs and KVIs at three levels—network, service, and societal. These indicators serve as the foundation for evaluating the technical and operational performance of 6G networks while also integrating broader socio-economic considerations. The methodology for KPI selection and validation draws upon insights from standardization bodies such as ITU, ETSI, and 3GPP, as well as findings from parallel 6G initiatives within the SNS Stream B and C projects and the 5GPPP Technology Monitoring and Validation Working Group (TMV WG).

The SAFE-6G framework introduces a user-centric approach, integrating cognitive coordination, AI-powered trust functions, and a robust validation methodology. These components are designed to enhance network intelligence, security, and reliability while ensuring alignment with societal values, including sustainability, digital inclusion, and privacy protection. The document details the methodologies employed for KPI measurement, validation, incorporating AI-driven monitoring mechanisms to optimize performance and trustworthiness dynamically.

The report also discusses the role of KVIs as complementary metrics to KPIs, ensuring that technological advancements in 6G align with broader societal priorities. These KVIs are linked to global frameworks such as the United Nations Sustainable Development Goals (SDGs) and the European Green Deal, emphasizing the importance of environmental sustainability, ethical Artificial Intelligence (AI) governance, and equitable digital access.

The validation methodology proposed in this deliverable is designed to be iterative, integrating real-world testing environments to assess SAFE-6G functionalities comprehensively. The approach encompasses multiple testing levels, including component-level verification, system-level evaluation, and large-scale deployment trials, ensuring the robustness and scalability of the framework.

The outcomes of this deliverable contribute to the overarching goal of the SAFE-6G project: developing a next-generation network framework that is not only technically advanced but also resilient, secure, and aligned with human-centric values. By establishing a rigorous methodology for KPI and KVI evaluation, this work lays the groundwork for future research, standardization efforts, and industrial applications in the evolving 6G landscape.

KEYWORDS

KPI, KVI, Cognitive Coordinator, 6G, Methodology, Trustworthiness, System Requirements, Stakeholder, Use case

TABLE OF CONTENTS

List of TABLES

List of FIGURES

1 INTRODUCTION - ABOUT THIS DOCUMENT

1.1 DELIVERABLE CONTEXT

This document presents the outcome of **Task 2.4**, focusing on the definition and validation methodology for Key Performance Indicators (KPIs) and Key Value Indicators (KVIs) essential for future user-centric 6G networks. The KPIs and KVIs have been carefully defined across three critical levels: network, service, and societal, ensuring comprehensive coverage of technical, operational, and socio-economic aspects relevant to 6G deployment. The requirements, KPIs and KVIs from the Use-Case perspective, were defined in Deliverable D2.3 (at Month 12).

The initial list of KPIs shown at Table 1, that was suggested in the project proposal, serves as the foundational basis for further development in this task. At the network level, these KPIs include capacity, speed, latency, user density, location accuracy, and energy efficiency. Service-level KPIs address infrastructure scalability, service creation time, and CAPEX/OPEX for network management. At the societal level, considerations such as Artificial Intelligence (AI) trustworthiness, data governance, and alignment with United Nations (UN) Sustainable Development Goals (SDGs) are key focus areas.

KPI Family	*KVI/KPI*	*Description*	*6G Impact*
Security	Level of Trustworthiness	The guaranteed level of trust/protection against certain threats and attacks.	More stringent due to pervasive utility of 6G and burgeoning risk level.
Response time	Time to respond	Time for SAFE-6G user-centric functions to counteract in case of malicious activity.	Much smaller due to compressed timescale of 6G networks, e.g., an attack can cause havoc at an order or faster.
Service availability and reliability	Coverage	The coverage of SAFE-6G user-centric functions over the 6G service elements and functions.	More challenging due to diverse 6G technologies and ultra-distributed functions.
AI	Atomicity and Cognitive Level	A measure of how autonomic and cognitive SAFE-6G AI/ML can act.	Expected to be easier to implement with pervasive AI, x(AI) and MLOps.
AI	AI robustness	The robustness of AI algorithms in the network hardened for trustworthiness.	It is more difficult to maintain consistency systemwide but more critical due to AI's role in 6G.
AI	AI model convergence time	Time for learning models working for security to converge.	Although more advances in AI/ML models are emerging and hardware capabilities are improving, the data availability and complexity are challenging factors for this KPI.

Response time	SAFE-6G user-centric Function Chain round-trip-time	Time for chained SAFE-6G user centric functions to process for ingest, analyze, decide and act (related to "Time to respond" KPI).	Trustworthiness architecture in 6G is supposed to be more distributed, leading to challenges. But at the same time, cognitive driven frameworks will help and need to be assessed.
Energy	Cost to deploy SAFE-6G user-centric functions.	Various cost metrics for measuring the cost of deployment for the five SAFE-6G user-centric functions.	Substantially increases due to complexity, thus in 6G it is expected to be more challenging to meet low target KPI values.

Table 1. DoA initial KPI list proposal.

Input from relevant standardization bodies (ITU, ETSI), 6G initiatives (SNS Stream B and C projects) and working groups (5GPPP TMV WG) has been incorporated to ensure that the defined KPIs and KVIs are aligned with the evolving standards and requirements. Additionally, the validation methodology is designed to be integrated into an AI/ML-driven SAFE-6G framework, emphasizing automation and reusability. The actual validation and measurement procedures will be executed in WP5.

1.2 THE RATIONALE BEHIND THE STRUCTURE

The structure of this deliverable is designed to comprehensively address the critical aspects of KPI and KVI definition and validation methodology for SAFE-6G. This includes the integration of insights from Deliverables D2.1 and D2.2, which laid foundational elements for user-centric frameworks and preliminary metrics identification. The organization of this document enables a logical flow, guiding the reader through the theoretical foundations, methodological advancements, and practical implementations.

Document Structure Overview:

1. **Introduction - About this document**: This section provides the deliverable's context, explaining its role in the broader SAFE-6G project, the rationale behind the document's structure, and a summary of the key outcomes. It establishes the background and significance of defining KPIs and KVIs for the SAFE-6G framework while connecting these efforts to previous deliverables and project objectives.

2. **SAFE-6G framework overview:** This section explains the overall SAFE-6G framework, focusing on its holistic design to address trustworthiness in future 6G networks. It highlights the roles of the cognitive coordinator and the five user-centric functions—Safety, Security, Privacy, Resilience, and Reliability—and describes how these components contribute to the framework's ability to meet the technical, operational, and societal demands of 6G.

3. **SAFE-6G KPIs & KVIs: Methodology and related state of art:** This section explores the current state of research and the implementation of KVIs, drawing from influential frameworks and

projects such as the 6G-AI White Paper and Hexa-X project. It also highlights how 6G can contribute to achieving the UN SDGs, reinforcing technology's role as a driver for global progress.

4. **SAFE-6G KPIs:** This section defines the KPIs developed for SAFE-6G, detailing their alignment with the project's trustworthiness objectives. It explains the KPI methodology, which incorporates technical, service, and societal dimensions, and presents a comprehensive list of indicators tailored to evaluate the framework's components, including the cognitive coordinator, trust functions, and supporting technologies.

5. **SAFE-6G KVs & KVIs:** This section introduces KVIs that integrate societal and environmental considerations into the evaluation framework. It emphasizes the connection to global frameworks such as the UN SDGs and the European Green Deal and explains how KVIs serve as a complementary measure to KPIs, ensuring that 6G development aligns with broader societal priorities.

6. **Validation methodology:** This section describes the methodology for validating KPIs and KVIs within the SAFE-6G framework. It outlines the components, testbeds, and approaches for measuring performance and societal impacts. SAFE-6G will follow a well-structured framework, as part of the DevOps approach, that facilitates several tests per each new service from the development, integration, and deployment phases ensuring on one hand that each new service addresses the requirements and KPIs and on the other hand that is compatible with the innovative features that SAFE-6G platform offers.

7. **Testing Tools:** This section identifies the tools and technologies selected to validate KPIs and KVIs across the SAFE-6G framework. It explains the mapping of specific tools to corresponding indicators, ensuring efficient and seamless validation. The section also highlights how these tools will integrate into the SAFE-6G platform, leveraging existing industry standards and innovative methodologies for comprehensive performance measurement.

8. **Conclusion:** This section summarizes the findings and implications of the deliverable, highlighting its contributions to SAFE-6G's vision of secure, trustworthy, and sustainable 6G networks. It also reflects on the broader impacts of the defined KPIs and KVIs, providing a foundation for future research and development in 6G technologies.

1.3 Outcomes of the Deliverable

The primary outcomes of this deliverable include:

1. **Comprehensive KPI and KVI Framework:**
    - Definition of KPIs across safety, security, privacy, resilience, and reliability.
    - Incorporation of societal-level KVIs to bridge technical objectives with societal values, aligned with SDGs and environmental goals.
2. **Methodological Advancements:**

- Development of a robust validation methodology integrating AI/ML tools for real-time KPI assessment.
- Introduction of the Level of Trustworthiness (LoTw) metric as a unified measure of system reliability and user-centric trust.

3. **Alignment with Future 6G Needs:**
   - Identification of enablers and blockers for KPI realization, ensuring readiness for complex 6G deployments.
   - Insights into the environmental, societal, and economic impacts of SAFE-6G, providing a roadmap for sustainability in 6G ecosystems.
4. **Practical Implementation Tools:**
   - Compilation of tools and technologies to validate KPIs and KVIs, integrated with the SAFE-6G platform.
   - Strategies to ensure scalability, resilience, and efficiency of SAFE-6G operations.

2 SAFE-6G FRAMEWORK OVERVIEW

The SAFE-6G architecture is designed around core building blocks that ensure a user-centric, trustworthy, and adaptable 6G ecosystem, as can be observed in Figure 1. It is structured as follows:

- **Chatbot & Cognitive Coordinator**: The AI-powered chatbot facilitates user interaction, enabling trust-based requests, which the Cognitive Coordinator interprets to determine the required LoTw and orchestrate trust functions accordingly.
- **Trust Functions (TFs)**: A set of AI-assisted mechanisms ensuring system trustworthiness across five key dimensions:
    - Safety
    - Security
    - Privacy
    - Resilience
    - Reliability

  Note: The three planes that are usually referred to are: Application plane with metaverse, 6G Core and Edge-cloud continuum
- **Metaverse application plane**
    - User-Centric Service/App Plane: Supports personalized applications with API-driven interactions.
- **6G System Planes**: SAFE-6G operates across three interrelated planes.
    - User-Centric System Openness Plane: Provides standardized API exposure via CAPIF for system programmability.
- **User-Centric Edge-Cloud Continuum Plane:**
    - Enables distributed deployment of applications and network functions for optimized user experience.
- **MLOps & DataOps:**
    - MLOps ensures efficient deployment, training, and maintenance of AI models across the framework.
    - DataOps provides automated processing of real and simulated data to support AI operations.
        - This architecture enables SAFE-6G to dynamically adapt to user-defined trust levels, ensuring a secure, resilient, and intelligent 6G environment.

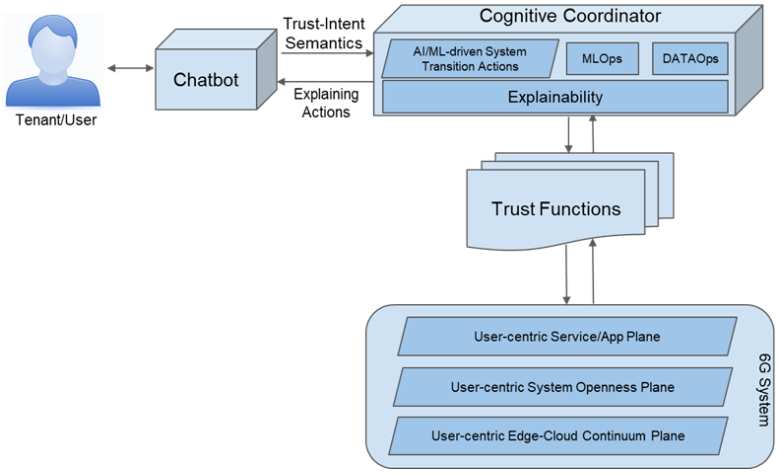

Figure 1. High-level view of SAFE-6G Reference Architecture

As 6G evolves, the need for user-centric, customized connectivity grows. SAFE-6G shifts from traditional operator-centric networks to a personalized, user-focused model, enabling secure, reliable, and adaptive services.

Unlike conventional architectures, SAFE-6G provides each user with a dedicated system instance, allowing individualized management of data, policies, and mobility, as well as the creation of custom VPNs and full control over data ownership.

The architecture features five Application Functions (AFs) acting as trust functions to ensure security, integrity, and tailored service delivery. By leveraging the edge-cloud continuum, it reduces latency, supports dynamic network slicing, and enhances 3GPP interfaces for an adaptive and human-centric 6G system.

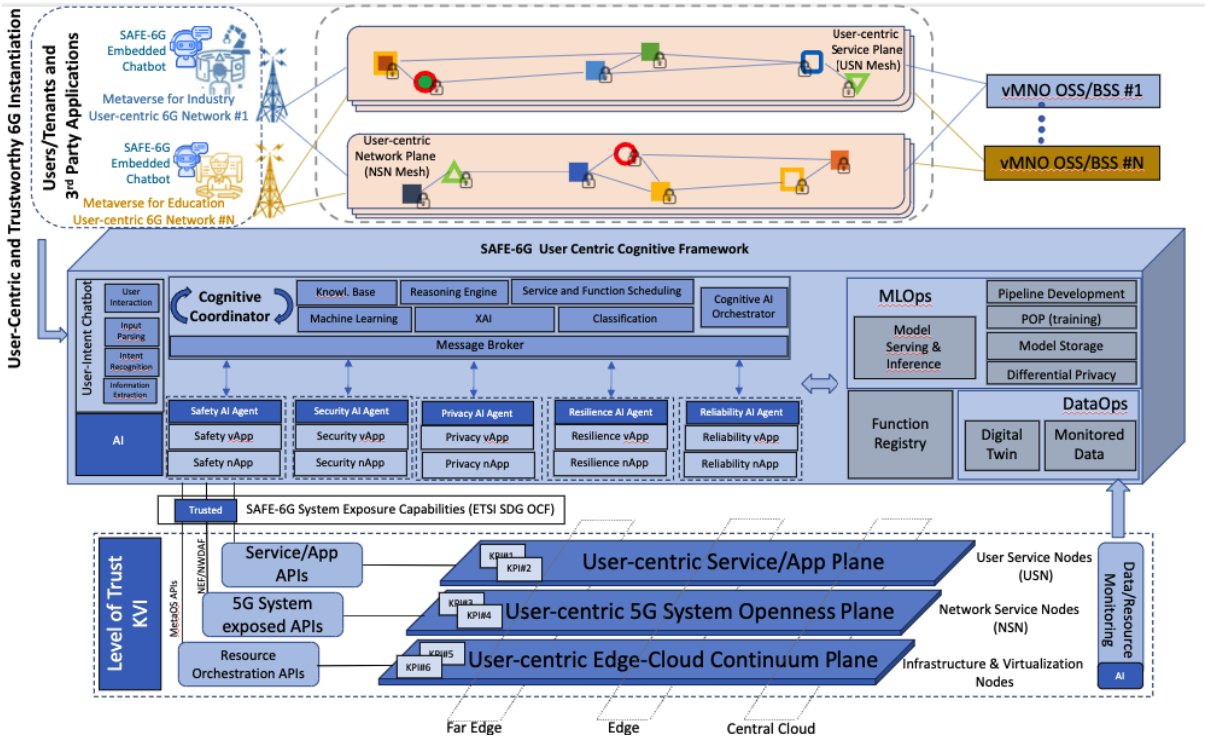

Figure 2. Functional View of SAFE-6G Reference Architecture

2.1 TRUST FUNCTIONS

SAFE-6G introduces five enablers/functions for trustworthiness, namely Safety, Security, Privacy, Resilience and Reliability, whose service placement and provision are coordinated by the Cognitive Coordinator in conjunction with the AI-driven resource and service mesh orchestrators, achieving an advanced and fully automated system (zero-touch) across all the network layers that form the SAFE-6G ecosystem.

Safety within SAFE-6G integrates the Software Defined Perimeter (SDP) paradigm within the User-Centric 6G Packet Core infrastructure. This technology establishes secure, individualized perimeters around critical network services such as nodes, controllers, and data centers. By doing so, SAFE-6G limits exposure to potential security threats and ensures that access is granted only to authorized users based on their specific needs. The SDP approach, grounded in the zero-trust security model, employs micro-segmentation of entitlements, creating a finely tuned security environment that adapts to the user's requirements. This ensures that the network remains resilient against unauthorized access and other security threats, safeguarding the infrastructure while providing a tailored experience for each user.

Security within SAFE-6G is further enhanced using blockchain-based verifiable credentials and the tokenization of user actions. These technologies ensure that all access and interactions within the network are tightly controlled and verified, reducing the risks associated with identity theft, data breaches, and other cyber threats. The decentralized nature of blockchain provides a transparent and tamper-resistant system, while regular security audits ensure that vulnerabilities are identified and

addressed proactively. By integrating these advanced security measures, SAFE-6G creates a robust and trustworthy environment, essential for handling the sensitive and personal data internal to the 6G ecosystem.

Privacy is a fundamental pillar of the SAFE-6G project, seamlessly integrated into its user-centric architecture. The project introduces mechanisms that allow users to define their privacy preferences through intuitive interfaces or APIs, depending on their role within the network. These privacy requirements are fed into the SAFE-6G Cognitive Coordinator, influencing decisions related to resource placement and management by AI-driven orchestrators. By embedding privacy considerations into the core operational processes, SAFE-6G ensures that user data is managed with the utmost confidentiality, aligning with the highest standards of user trust and control.

Resilience within the SAFE-6G network is achieved through an intent-based approach that utilizes a sophisticated ontology of intent, built on knowledge management and semantic modelling techniques. The resilient function is designed to interpret user intents and translate them into actionable network operations that optimize service delivery. Machine learning (ML) and a knowledge base structured using the RDF Schema enable the network to adapt dynamically to changing conditions and user needs. This adaptability ensures that the network remains robust and capable of delivering continuous, efficient service, even under varying operational demands.

Reliability of the SAFE-6G system is underpinned by comprehensive service and reliability profiling, which involves multi-layer data collection across physical, virtual, and application layers. This data feeds into federated learning-based mechanisms that ensure data privacy while enabling advanced machine learning capabilities. By continuously monitoring service performance and simulating various operational conditions, SAFE-6G can detect and respond to potential threats in real time. This approach ensures that the network maintains high reliability and consistent service quality, even in the face of diverse challenges such as DDoS attacks or other forms of intrusion.

2.2 COGNITIVE COORDINATOR

The SAFE-6G Cognitive Coordinator is an advanced intent-handling component that plays a key role in managing the trustworthiness of the SAFE-6G user centric cognitive framework. It interprets user-defined trust requirements conveyed via the AI Chatbot, calculates the desired LoTw and coordinates the system's transition to a trustworthy state. This is achieved by orchestrating the five SAFE-6G trust functions: Safety, Security, Privacy, Resilience, and Reliability.

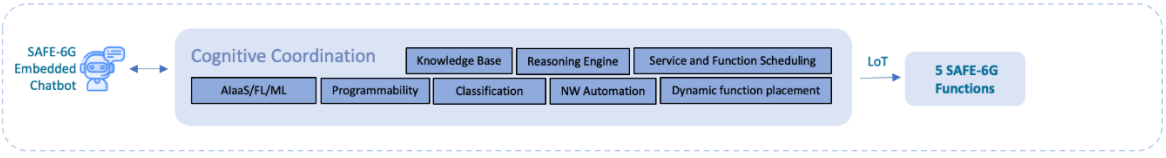

Figure 3. SAFE-6G High level view of Cognitive Coordinator

The Cognitive Coordination process involves managing and orchestrating the classification and reasoning activities to ensure optimal and achievable LoTw in the SAFE-6G framework. Initially, the classification component produces a non-calibrated level of trustworthiness (nLoT) by extracting features, processing data, and running an AI algorithm on the input. The Reasoning Engine then refines this initial nLoT, profiling each trust function and querying a knowledge base to integrate relevant historical data and system metrics. Each trust function sets its boundaries based on real-time metrics, such as resource usage and network conditions, defining feasible levels of trustworthiness. Finally, the Cognitive AI Orchestrator then combines all the outputs of the classification and the reasoning engine components, along with system metrics, network data, and contextual information, to produce a calibrated Level of Trustworthiness (cLoT). This cLoT is retrieved by the corresponding AI agent of each trust function, which interacts with the local knowledge base and policy to determine necessary actions.

2.3 MLOps/DataOps

The MLOps framework supports and eases the development and execution of AI/ML algorithms across multiple domains. The framework will enable the management of a large part of the lifecycle of AI/ML models, from design, training, and evaluation to deployment in production environments at different domains (far/extreme edge, edge, cloud), allowing their smooth integration into the Cognitive framework. Therefore, SAFE-6G will cope with the challenge of having many closed control loops at every segment of the network, facilitating the AI/ML model development operations in architecture. The MLOps Framework serving infrastructures will allow the serving of multiple models simultaneously, managing both versioning, and model labelling and separate testing and production environments.

The DataOps component will support MLOps, with dedicated storage types and databases to support the AI/ML training processes. Part of this module is also the Network's Digital Twin (NDT), which extends the notion of digital twins in the domain of networking. In SAFE-6G, NDT will comprise three layers: (i) The first layer will be the physical network layer which is the target real world 6G network operating environment. (ii) The second will be the 6G twin layer which is the core of the NDT, and which integrates the data domain, the model domain, and the management domain. (iii) The third will be the application layer, which enables the two-way interaction between them as it provides the means to exploit the capabilities offered by the 6G twin layer. The entire process will complement the MLOps system of SAFE-6G, which is powered by AI/ML functions that are key to achieving the cognitive coordination autonomous feedback loop between the 6G physical network and the SAFE-6G framework. The AI/ML functions produced by the NDT and feed to the MLOps for initiative the cycle of continuous training and enhancement of their predictions will provide a powerful set of tools for analyzing and understanding complex data. In an NDT, there are typically large amounts of data generated by sensors and other sources, and ML techniques will help to extract insights from this data that would be difficult or impossible to obtain using traditional methods. Therefore, NDT in SAFE-6G will support data generation by simulation for the AI/ML training and inference processes.

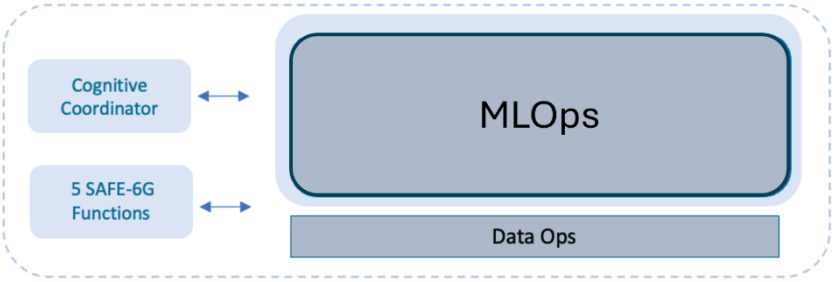

Figure 4. MLOPs/DataOps domain

2.4 EOSC ALIGNMENT IN SAFE-6G ARCHITECTURE

SAFE-6G embeds FAIR (Findable, Accessible, Interoperable, Reusable) principles and aligns with the European Open Science Cloud (EOSC) vision through its DataOps layer and observability components. This integration ensures that all KPI/KVI measurements, AI/ML datasets, and system logs generated by the Cognitive Coordinator, and Trust Functions are curated and published according to European Open Science standards.

Integration Points:

- **DataOps Layer as Data Fabric**:
  Acts as the central hub for data collection, normalization, and enrichment. All datasets are stored in standardized formats (CSV, JSON, RDF) and annotated with rich metadata.
- **Persistent Identifiers and Versioning**:
  Each AI/ML dataset is assigned to a DOI and version number to ensure traceability and reproducibility across SAFE-6G releases.
- **EOSC-Compliant Deposition**:
  SAFE-6G publishes datasets in trusted EOSC repositories such as Zenodo, ensuring open access for non-sensitive data and controlled access for sensitive data via EOSC AAI mechanisms. Examples of published datasets include:
  - Cognitive Coordinator Dataset of User Intent for Trustworthiness on five principles (Reliability, Privacy, Security, Resilience, and Safety) [8]
  - An Open-Source Framework and Dataset for Multi-Layer Monitoring and Predictive Autoscaling of 6G Video Streaming [9]
- **Interoperability and Reusability**:
  All datasets include provenance information, licensing details, and links to related SAFE-6G deliverables. This guarantees interoperability with other EOSC services and supports cross-project benchmarking.

2.5 CHATBOT

The Chatbot domain within the SAFE-6G framework is the first domain that contributes to the enhancement of user interaction and ensuring trustworthiness in the User-centric Distributed 6G Core.

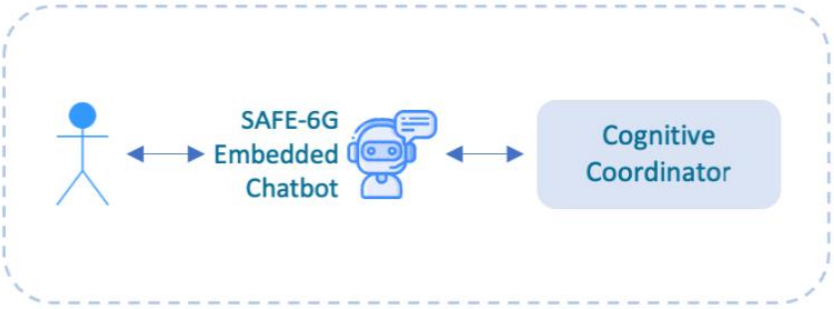

Figure 5. High-level view of the chatbot

More specifically, it will allow users to interact with the system to request the LoTw needed through the SAFE-6G Cognitive Coordinator, which extracts the LoTw across the five SAFE-6G user-centric functions: Safety, Security, Privacy, Resilience, and Reliability. The procedure starts by prompting service-driven questions to the user. Then, this acts as an input and is processed by the NLP component that converts the raw text into structured data that is later an input to the Cognitive Coordinator Component. A virtual chatbot window in XR will be present in the metaverse Use Case applications. This way, users will be able to request the LoTw they need, which is transmitted to the SAFE-6G chatbot application through the metaverse manager component. Users will also receive explainability (XAI) responses on the AI models that are going to be applied to achieve the LoTw.

2.6 TRACEABILITY TO REQUIREMENTS

To ensure consistency and traceability throughout the project's development and validation phases, the SAFE-6G framework components outlined in this deliverable (D2.4) are directly mapped to the technical requirements defined in D2.1. This explicit link is essential for anchoring the Key Performance Indicators (KPIs) and Key Value Indicators (KVIs) to the core design specifications of the architecture.

SAFE-6G Components	D2.1 Section/Requirements category
Chatbot Requirements	Section 5.2
Cognitive Coordinator (CoCo)	Section 5.3
Safety Function	Section 5.4
Security Function	Section 5.5
Privacy Function	Section 5.6
Resilience Function	Section 5.7
Reliability Function	Section 5.8
MLOPS Framework	Section 5.9
Core requirements	Section 5.10
Continuum Requirements	Section 5.11

3 SAFE-6G KPIs & KVIs: METHODOLOGY AND RELATED STATE OF ART

3.1 STATE OF THE ART FOR KVs AND KVIs

This section explores the current state of research and the implementation of KVIs, drawing from influential frameworks and projects such as the 6G-AI White Paper and Hexa-X project. It also highlights how 6G can contribute to achieving the UN SDGs, reinforcing technology's role as a driver for global progress. By embedding values such as trust, inclusiveness, environmental responsibility, and economic equity into network design, the SAFE-6G initiative sets the stage for a more human-centered, sustainable digital future.

3.1.1 UNITED NATIONS SUSTAINABLE DEVELOPMENT GOALS AND MOBILE INDUSTRY

In 2015, the UN defined the 17 SDGs [1] and the targets that should be achieved by 2030, inviting governments, industrial sectors, and stakeholders to collaborate in accomplishing these goals. Given that the commercial debut of 6G is scheduled for 2030, this developing technology should contribute to the achievement of the SDG targets.

According to GSMA reports, the mobile industry has made beneficial contributions to the achievement of these goals, but recent global events have been a major setback for the SDGs [2] . The GSMA assesses the influence of the mobile industry on the achievement of the SDGs through studying aspects such as mobile coverage, technology adoption, and the level of technology consumption in the proportion of the world's population, as well as the positive or negative impact of MNO operations on the environment and society. Each of these assessed variables has an impact at various levels on each of the 17 SDGs.

The Finnish 6G Flagship [3] research program has published a white paper that aims to establish a clear link between SDGs, existing mobile industry KPIs, and 6G. The Figure 6 represents a high-level interaction between the SDGs and the envisioned 6G capabilities, including networking, processing, actuation, sensing, data, and intelligence everywhere. 6G aspires to empower individuals through innovative services, sense/monitor the environment on a local and global scale and strengthen the globe by contributing to the SDGs.

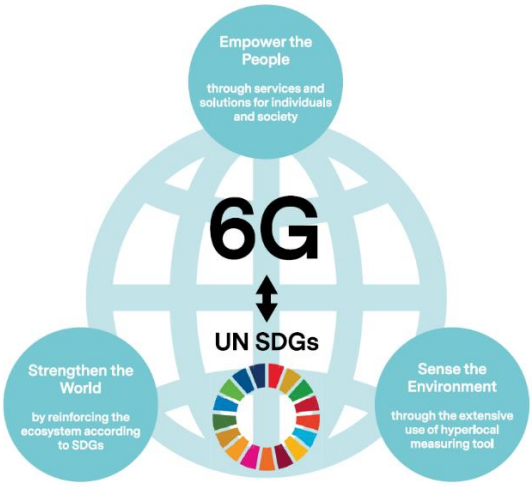

Figure 6. Three pillars connect 6G with the UN SDGs[1]

The 6G Flagship white paper examines the current connection of the UN SDG with the mobile communications/ICT industry, as well as the interaction with existing indicators from the UN SDG framework, GMSA, and ITU. Going a step further, the white paper identifies which specific SDG targets may be impacted by 6G. The authors discovered a connection between a set of specified aims and UN metrics and then focused on how 6G can contribute to increasing these indicators.

3.1.2 6G-AI WHITE PAPER ON KVS AND KVIS

The 6G AI whitepaper [1] created by the 6G Infrastructure Association Vision and Societal Challenges Working Group, Societal Needs and Value Creation Sub-Group, defines value as what is essential to people and society and can be addressed (partially) or indirectly by future network technology knowing value thus entails knowing all the factors (e.g., context, circumstance, ecosystem dynamics) to cooperatively identify the greater good with stakeholders. According to the WG, this requires a mentality shift and reframing, with innovation driven not only by technology specifications and economic goals, but also by a greater awareness of the ramifications for society.

Key Value Indicators (KVIs) are designed to supplement a more traditional performance-based (KPI) approach by (a) demonstrating the validity of innovation in satisfying societal needs, and (b) influencing innovation trajectories to benefit values (apart from user needs). It is used to set measurable goals (quantitatively, the number of people a service serves, or qualitatively, the impression of value fulfilment when utilizing a service). The end goal is to create a process that empirically demonstrates how an invention adds value. They are described together with enablers that help reach the targets.

Drawing from the societal values highlighted in the UN SDGs, the European Green Deal, human-centered design principles, and the broader European ideals of strategic autonomy and technological

[1] Figure originally taken from the White Paper:
https://www.researchgate.net/publication/341069017_White_Paper_on_6G_Drivers_and_the_UN_SDGs

sovereignty, the WG proposes Key Value Themes. These themes can serve as the foundation for developing indicators based on 5G/6G use cases.

- Environmental sustainability
- Social sustainability
- Economic sustainability
- Democracy
- Cultural connection
- Knowledge
- Privacy and confidentiality
- Simplified life
- Personal freedom
- Personal health and protection from harm
- Trust

Also, the white paper proposes that any produced indicator should support assessment path which may comprise various combinations of the following:

- **Objective evaluation by experts and representatives:** This includes expert assessments and measurements taken directly from the deployed networks.
- **Subjective evaluation by representatives:** This involves data gathered from trials, experiments, interviews, questionnaires, and focus groups.

It also describes a four-step procedure for developing KVIs:

- **Define key values (KVs):** Determine the appropriate KVs for the use case by understanding the societal difficulties or "pain points" that motivate these values.
- **Identify the KVI:** Estimate the magnitude of the effect by determining what percentage of the population is realistically affected by the use case.
- **Determine enablers and blockers:** Identify technological and societal elements that may help or impede the use case.
- **Quantify indicators with KPIs:** When possible, quantify the indicators using KPIs. These can be technical or numerical targets that, according to expert stakeholder assessment, enable the KVs. Each KPI should be clearly described to show how and to what extent it represents an improvement towards the KVI.

And emphasizes essential critical points:

- **Project and use case goals must be clarified prior to formulation metrics:** These goals should extend beyond technical achievements to include an understanding of the context in which the technology is applied and how it is utilized. This necessitates input from a diverse range of stakeholders relevant to those contexts, such as social scientists, natural scientists, practitioners, community representatives, and local businesses.

- **The formulation, assessment, and causality of KVIs should involve relevant experts and stakeholders.** This is necessary because the causal chains from measures to impact are complex and often extend beyond project timeframes. Involving diverse experts helps avoid the unintentional inflation of value terms or 'value-washing' (like green-washing or ethics-washing).
- **KVIs** serve as demonstrations of **potential value** rather than providing directions or instructions for designing impactful technology.

3.1.3 HEXA-X

HEXA-X [5] tries to represent the societal value provided by 6G by mapping the project use cases to global UN SDG targets, as well as extra societal targets derived from the investigated use cases that are not derived from SDGs. As a result, it became necessary to define a set of KVIs for each target that should be taken into consideration when creating 6G technology. Given that the SDGs and other targets currently specified are not straight away applicable to technological development, HEXA-X intends to define specific KVI that are directly applicable to 6G development. They estimate that some KVIs will be able to be examined directly, while others will be assessed indirectly using KVIs.

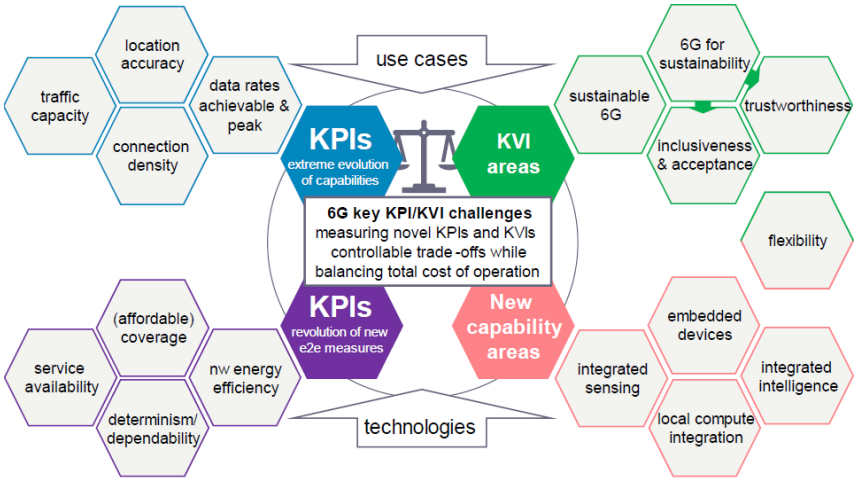

Figure 7. Clustering of KPIs and KVIs

The HEXA-X project undertook an early assessment of general social values with the aim of developing an initial set of KVIs exclusive to 6G technology. As indicated in the previous paragraph, these efforts expanded on the UN SDGs while also looking more broadly at other policies that define social, economic, and environmental values. It stated that, while some KVIs can be examined right away, the majority are appraised using proxies based on a combination of both qualitative and quantitative data. It identified critical values for the Hexa-X project (and thus features of beyond 5G and 6G).

- Sustainable 6G,
- 6G for sustainability,
- Inclusiveness and Acceptance,
- Trustworthiness,
- Flexibility

Sustainability in both directions in the list above addresses environmental, social, and economic aspects, as well as how to meet current needs without affecting future generations' ability to meet their own. This involves keeping in mind the influence of the solutions' lifecycle (e.g., 5G and 6G technology and practices) on sustainability (e.g., their effects on the environment, human rights, and socioeconomic progress). It is also essential to explore how to build the solutions as enablers of sustainability throughout time.

3.2 SAFE-6G KVs & KVIs: METHODOLOGY AND ANALYSIS

Given the current state of art, SAFE-6G's main premise is the selection of KVIs that are relevant for the scope of the project, on the basic principle that "6G networks can positively impact societies." This vision is grounded in the belief that the development and deployment of 6G technology must be driven by values that prioritize sustainability in its various forms. As such, the technology's design framework needs to ensure that environmental, societal, and economic factors are central considerations in the decision-making process:

Sustainability in SAFE-6G is defined across three critical dimensions:

- The **environmental** axis focuses on minimizing energy consumption, reducing CO2 emissions, and managing water usage. These environmental concerns are paramount in ensuring that future 6G networks are aligned with global efforts to mitigate climate change and preserve natural resources.
- The **societal** axis emphasizes the importance of upholding democratic values, promoting digital inclusion, and expanding access to knowledge. This reflects the role that 6G can play in fostering more inclusive and equitable societies, where all individuals can benefit from the opportunities provided by advanced technology.
- The **economic** axis targets the cost-efficiency of living and working, particularly in rural and underserved areas. Through 6G, remote work and digital services can be enhanced, driving economic sustainability by reducing costs and improving quality of life in these regions.

The potential of the ICT industry to drive sustainability is understood in two distinct ways.

First, there is the direct impact, referred to as "Sustainable ICT." This involves developing technologies that can be measured and quantified in terms of their sustainability contributions. The overarching goal here is to create a sustainable 6G infrastructure, where improvements in energy efficiency, resource management, and operational costs are key outcomes. For example, one direct impact might be seen in the significant reduction of energy consumption in data transmission through more efficient network designs.

Second, there is the indirect impact, which is framed as "ICT for sustainability." This focuses on how 6G networks can serve as enablers for other industries and verticals to enhance their own sustainability practices. The goal is to position 6G as an enabler of sustainability across various sectors, providing tools and capabilities that allow industries to make more informed and sustainable

decisions. For instance, smart city technologies powered by 6G could help urban centers optimize resource usage, from energy to water, significantly reducing their environmental footprint.

Through this dual approach, ensuring both that the ICT sector itself is sustainable and that it empowers other sectors to adopt sustainable practices, SAFE-6G aims to drive meaningful and measurable progress towards a more sustainable future. The specific KPIs and KVIs selected for this project reflect this approach, focusing on the balance between maintaining technological performance and promoting broad societal benefits.

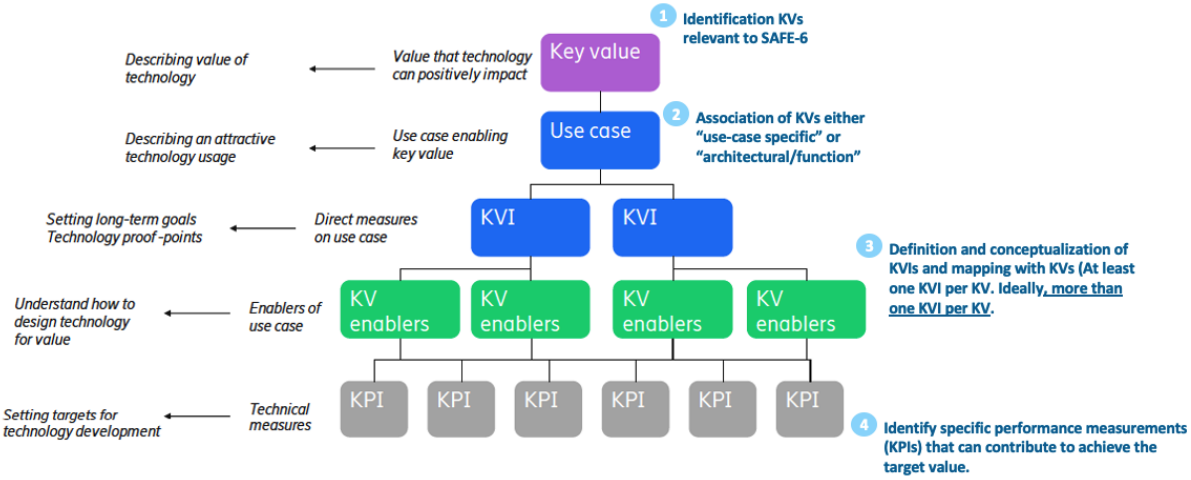

Figure 8. SAFE-6G KVI procedure structure based on reference [6] .

KVIs represent a transformative approach to evaluating innovation within 5G and 6G networks, shifting the focus from purely technical and economic outcomes to the broader societal and environmental values they serve. Based on the foundational insights from the 6G Infrastructure Association's whitepaper detailed in Section 3.1.2, KVs emphasize critical pillars such as:

- **Trust**
- **Security, Privacy and Confidentiality**
- **Economic Sustainability and Innovation**
- **Environmental Sustainability**
- **Societal Sustainability**
- **Simplified Life**
- **Digital Inclusion**

These indicators address the essential needs of people and society, aligning with global initiatives like the UN SDG and the European Green Deal. By adopting a cooperative and inclusive methodology that considers the contextual, societal, and ecological dynamics of innovation, KVIs complement traditional KPIs.

This reframing emphasizes measurable contributions to societal challenges, ensuring that innovation is not only efficient but also meaningful. The development of KVIs involves defining societal values, identifying relevant use cases, determining enablers and barriers, and quantifying impacts through KPIs.

Key Value (KV)	***Key Value Indicators (KVI)***
Trust	Trustworthiness
Security, Privacy, and confidentiality	Security and Privacy
Environmental Sustainability	Sustainable Resource Utilization
Environmental Sustainability	Sustainable Cities and Communities
Simplified life	Dynamic Programmability and Determinism Indicator
Digital inclusion	Global Scalability and Affordability
Societal Sustainability	Data Protection Index

Table 2. SAFE 6G KVs vs KVI reference

3.3 SAFE-6G KPI METHODOLOGY

The identification of KPIs for the SAFE-6G project follows a structured and comprehensive simple methodology. This process is centered around a careful review of the key elements that ensure the Security, Safety, Privacy, Resilience, and Reliability of the 6G network.

The SAFE-6G Framework is the foundation of this methodology, designed to build trustworthy user-centric 6G networks. Its core functions, Security, Safety, Privacy, Resilience, and Reliability, serve as the primary reference points. Each of these functions is used to define specific KPI areas, ensuring that the indicators align with the overall objectives of creating a network that can meet future demands for performance, protection, and trustworthiness.

An analysis of the core functions within the SAFE-6G framework and the technology founding them is essential in identifying relevant KPIs. **Security**-related KPIs focus on protecting data, ensuring robust authentication, and maintaining network integrity. **Safety** metrics focus on the secure and reliable operation of critical services. **Privacy** indicators highlight the importance of data governance and user privacy. **Resilience** KPIs assesses how well the network can recover from disruptions and maintain service continuity throughout the cloud continuum. **Reliability** metrics emphasize the consistent and dependable delivery of services.

Apart from key technology, Standard Developing Organizations (SDO) such as ITU, ETSI, and 3GPP are also evaluated. These standards provide the necessary guidelines and requirements for defining technical KPIs, ensuring that they are consistent with international benchmarks. The methodology integrates these standards to ensure compliance and forward-looking relevance for future 6G deployments. Special focus has been set on the TMV [1] WG.

Additionally, the methodology draws on insights from other finalized and current ICT and H2020 projects, including key learnings from ongoing and past initiatives under the Horizon 2020 program. These projects provide valuable references and benchmarks for KPI definition, ensuring that the indicators are informed by real-world results and experiences. By leveraging knowledge from these projects, the methodology incorporates cutting-edge research and development insights to ensure KPIs are both practical and innovative.

3.3.1 KPI MEASUREMENT ARCHITECTURE AND REFERENCE POINTS

To ensure methodological rigor, KPI collection in SAFE-6G follows a defined measurement architecture spanning three layers of the system:

- Network Layer (Core and Edge): Measurements such as latency, throughput, jitter, and packet loss are collected at 3GPP reference points (e.g., between UPF–gNB) using standard test protocols (iPerf3 for throughput/latency, SNMP for operational metrics).
- Service Layer: Metrics like service uptime, orchestration time, and API responsiveness are measured at the aerOS meta-OS interfaces using Prometheus exporters and Cilium/Hubble telemetry.

- Societal Layer: Data for Key Value Indicators (e.g., sustainability, inclusion) are aggregated through the DataOps framework, correlating KPI-level metrics with environmental or user-level data.

All metrics are collected via the SAFE-6G Data Fabric, ensuring synchronization and timestamp alignment across distributed nodes. Test executions are orchestrated through CI/CD pipelines, guaranteeing repeatability and traceability.

Finally, the identification process is iterative and collaborative, involving inputs from consortium members, industry experts, and standardization bodies. This collaborative approach ensures that the KPIs are measurable and aligned with broader societal objectives, such as the UN SDGs. Stakeholder feedback is incorporated to refine and adjust the KPIs throughout the process, ensuring they are relevant for both technical implementation and societal impact.

This methodology, based on the thorough examination of the SAFE-6G framework, its core functions, relevant standards, and lessons learned from related projects, ensures that the KPIs are both comprehensive and future-proof, ready to guide the development of trustworthy 6G networks.

3.3.2 KPI Measurement Specification

All KPIs proposed for assessment of SAFE-6G components and use cases follow same structure described next:

- **KPI category**: referring to the targeted KPIs TMV WG defined families (for example, Data Rates and Capacity, Latency, Mobility, Reliability and Availability, or other categories),
- **KPI Name**: given in the context of the project,
- **KPIs Definition**: detailed to ensure understanding of the true nature of the KPI,
- **Use case**: context of the KPIs in SAFE-6G project,
- **Data sources**: information on where and how the KPIs will be measured, referring to the layer, component, or the reference points where measurements will be collected,
- **KPI evaluation**: relative or absolute evaluation of the KPIs,
- **Verification method:** specific formula and/or procedure to calculate KPI.

Given the broad and ambitious scope of the KPI framework, SAFE-6G introduces a prioritization scheme that distinguishes between *core* and *extended* indicators across all Trust Functions (TFs). **Core KPIs** represent the essential, minimal set of metrics required to validate the operational effectiveness, responsiveness, and availability of each Trust Function within the SAFE-6G framework. They are directly linked to demonstrating compliance with project objectives and system-level performance targets. In contrast, **extended KPIs** provide complementary or higher-level insights—covering aspects such as optimization quality, user satisfaction, or orchestration efficiency—that enrich the analysis and support advanced validation, benchmarking, or future research. This distinction enhances practicality and focuses during testing and validation activities, ensuring that baseline evaluations remain achievable while still enabling deeper exploration of trustworthiness dimensions as the SAFE-6G platform matures.

4 SAFE-6G KPIs

The SAFE-6G project is committed to ensuring that its trustworthiness framework does not negatively impact the overall performance of the 6G system. To achieve this, SAFE-6G will evaluate the framework's influence on network performance, focusing on both the deployment phase and the session runtime of its five user-centric functions. This evaluation process will involve a combination of KPIs.

These indicators, as outlined in 5GPPP's Technology Monitoring and Validation Working Group (TMV WG) and the Strategic Research and Innovation Agenda 2021-2027, provide a comprehensive baseline for measuring performance across various network aspects. In addition to these widely recognized KPIs, SAFE-6G has introduced its own set of indicators, specifically focusing on 6G trustworthiness and security. The final KPIs selected for the project ensure that while the SAFE-6G framework enhances security, privacy, safety, resilience, and reliability, it does not compromise key network metrics such as latency, bandwidth, or overall system efficiency.

The section will explain in detail the full list of KPIs used within SAFE-6G, focusing on how each indicator is applied and measured to maintain optimal 6G performance alongside the advanced trustworthiness framework.

SAFE-6G project has identified the following lists of technical KPIs which have been clustered by TMV group KPI families:

- SAFE-6G general KPIs, which include:
  - Latency
  - Capacity
  - Energy
  - Operational network
  - Other
- SAFE-6G framework-specific KPIs:
  - Cognitive Coordinator
  - Trust Functions
  - Orchestrator
  - MLOps framework
  - Chatbot

4.1 SAFE-6G Framework KPIs

4.1.1 SAFE-6G Cognitive Coordinator KPIs

KPI category	*KPI name*	*KPI definition*	*KPI target value*	*Use case*	*Data sources*	*KPI evaluation*	*Verification method/tools*
Cognitive Coordinator	Mean Absolute Error (MAE)	The average of absolute differences between predicted and actual values. Measures overall prediction error considering a range [0,1].	≤ 0.20	Both UCs	SAFE-6G chatbot output data, representing user needs and intents.	Absolute	Use scikit-learns mean_absolute_error function or calculate manually.
Cognitive Coordinator	Mean Squared Error (MSE)	The average of squared differences between predicted and actual values. Penalizes larger errors more heavily.	≤ 0.15	Both UCs	SAFE-6G chatbot output data, representing user needs and intents.	Absolute	Use scikit-Learn's mean_squared_error function or calculate manually.
Cognitive Coordinator	Root Mean Squared Error (RMSE)	The square root of the average squared differences between predicted and actual values.	≤ 0.25	Both UCs	SAFE-6G chatbot output data, representing user needs and intents.	Absolute	Use scikit-learn's mean_squared_error and take the square root.
Cognitive Coordinator	R^2 Score (Coefficient of Determination)	Measures how well the predictions match the actual values. Ranges from 0 to 1, with higher values being better.	≥ 0.68	Both UCs	SAFE-6G chatbot output data, representing user needs and intents.	Relative	Use scikit-learn's r2_score function to evaluate.
Cognitive Coordinator	Mean Absolute Percentage Error (MAPE)	The average percentage error between predicted and actual values.	≤ 32%	Both UCs	SAFE-6G chatbot output data, representing user needs and intents.	Relative	Calculate manually during evaluation phase
Cognitive Coordinator	Provision of Level of Trustworthiness	The time required for the Level of Trustworthiness to be calculated by the Cognitive Coordinator.	< = 5 min	Both UCs	Various internal metrics	Absolute Evaluation	Time elapsed

Table 3. SAFE-6G Cognitive Coordinator KPIs.

The performed by utilizing sklearn library.

4.1.2 SAFE-6G TRUST FUNCTIONS' KPIS

4.1.2.1 SAFE-6G PRIVACY FUNCTION KPIS

KPI category	*KPI name*	*KPI definition*	*KPI target value*	*Use case*	*Data sources*	*KPI evaluation*	*Verification method/tools*
Latency	Privacy Score Calculation Latency	Measure the time taken from data input (service and CNC) to the calculation of the updated Privacy Score.	≤ 3 seconds	Ensuring timely updates to the Privacy Score.	System logs, performance monitoring tools.	**Absolute** evaluation of calculation latency.	Performance testing and log analysis.
Latency	Privacy Action Recommendation Latency	Measure the time taken to process data from the moment it is received to the point where a privacy recommendation is made.	≤ 3 seconds	Ensuring timely privacy recommendations in near real-time scenarios.	Network traffic logs, DSS processing logs.	**Absolute** evaluation of the time taken to process data.	Performance testing and log analysis during peak traffic periods.
Reliability and Availability	Privacy Function Uptime	Measure the percentage of time the Privacy Function is operational and available.	≥ 99.9%	Ensuring high availability of the DSS for continuous operation.	System uptime logs, monitoring tools.	**Absolute** evaluation of system uptime.	Monitoring system uptime logs over a specified period.
Performance	Privacy vApp flavor selection accuracy	Measure the accuracy on the vApp flavor selection from the AI Agent	≥ 80% satisfaction	Ensuring the Privacy Action suggestions meet expectations.	Function Logs.	**Relative** evaluation based on manual audits and log analysis.	Conducting periodic manual audits and log analysis.

Table 4. Privacy function Core KPIs.

KPI category	*KPI name*	*KPI definition*	*KPI target value*	*Use case*	*Data sources*	*KPI evaluation*	*Verification method/tools*
Latency	Privacy Function latency	The time required for the Level of Privacy to be provided by the Privacy Trust Function to the Cognitive Coordinator.	< = 3 min	Both UCs	Requested LoTw	Absolute Evaluation	Time elapsed

Table 5. Privacy function extended KPIs.

4.1.2.2 SAFE-6G SECURITY FUNCTION KPIS

KPI category	*KPI name*	*KPI definition*	*KPI target value*	*Use case*	*Data sources*	*KPI evaluation*	*Verification method/tools*
Blockchain	Peer Count	The total number of active peers in the blockchain network	>= 5 Peer	Both UCs	Network configuration and management tools	Absolute	Number of networks per in X period.
Latency	Transaction Latency	The time taken for a transaction to be confirmed and included in the blockchain.	<2 seconds for transaction endorsement, <2-3 seconds for final confirmation.	Both UCs	Logs from blockchain peers, orderers and chaincodes	Absolute	Time the peers require to send a new transaction to the network.
Blockchain	Block Finalization	The average time it takes to create a new block	< 5s	Both UCs	Logs from blockchain peers, orderers and chaincodes	Absolute	Time the network requires to add a new block to the ledger.
Blockchain	Transaction Throughput	Number of transactions processed per second (TPS)	100 - 200 TPS	Both UCs	Logs from blockchain peers, orderers and chaincodes	Absolute	Average transaction pers second in period.
Blockchain	Network Uptime	Percentage of uptime the blockchain network is up and available	> 99 %	Both UCs	Logs from blockchain network peers	Absolute	Consecutive uptime of network in period.
Security Trust Function	Provision of Level of Security	The time required for the Level of Security to be provided by the TF to the Cognitive Coordinator.	< = 3 min	Both UCs	Internal Metrics	Absolute Evaluation	Time elapsed

Table 6. SAFE-6G Security function Core KPIs.

KPI category	*KPI name*	*KPI definition*	*KPI target value*	*Use case*	*Data sources*	*KPI evaluation*	*Verification method/tools*
Blockchain	Smart Contract Deployment Success Rate	Percentage of successful smart contract (chaincode) deployments.	> 99 %	Both UCs	Logs from blockchain peers, orderers and chaincodes	Absolute	Consecutive successful smart contract deployments
Blockchain	Smart Contract Execution Success Rate	Percentage of successful executions of smart contracts.	> 95 %	Both UCs	Logs from blockchain peers, orderers and chaincodes	Absolute	Consecutive successful smart contract invocations
Blockchain	Smart Contract Response Time	Average response time for executing smart contracts.	< 2 second	Both UCs	Logs from blockchain peers, orderers and chaincodes	Absolute	Duration of smart contract invocation
DID Management	DID Generation Time	Local generation time for a unique identifier (DID) on a wallet.	< 100 ms	UC agnostic	Logs from device where secret key is stored (wallet)	Absolute	DID Creation logs
DID Management	DID Verification Percentage	Percentage of successful DID verification (cryptographically, integrity and governance).	> 99%	UC agnostic	DID Registry Logs (Blockchain) and Logs from device (wallet)	Absolute	DID Registry and Device Logs, Validation method of DID Document
Credential Verification	Credential Verification Percentage	Percentage of successful credential verification (cryptographically, integrity and governance).	> 99%	UC agnostic	DID Registry Logs (Blockchain) and Logs from device (wallet)	Absolute	DID Registry and Device Logs, Validation method of DID Document
Credential Management	Credential Generation Time	Time of credential generation on a wallet.	< 300 ms	UC agnostic	Logs from device where Credential is generated.	Absolute	DID Registry, Device and Agent Logs
Credential Lifecycle Management	Lifecycle List Creation	Duration of operation on credential lifecycle list.	<2-3 seconds for final confirmation	UC agnostic	DID Registry Logs (Blockchain) and Logs from device (wallet)	Absolute	DID Registry, Device and Agent Logs

Table 7. SAFE-6G Security function Extended KPIs.

4.1.2.3 SAFE-6G SAFETY FUNCTION KPIS

KPI category	*KPI name*	*KPI definition*	*KPI target value*	*Use case*	*Data sources*	*KPI evaluation*	*Verification method/tools*
Capacity	Tunnel Throughput Efficiency	The ratio of the actual occupied throughput to the QoS target throughput as defined by the service level agreements (SLAs)	>85%	Both UCs	SDPGW	Relative	SNMP features on the GW
Safety Trust Function	Provision of Level of Safety	The time required for the Level of Safety to be provided by the Tf to the Cognitive Coordinator.	< = 3 min	Both UCs	Internal Metrics	Absolute Evaluation	Timelapse
Scalability	Safety Service Resource requirement	Resources required for safety service that will be used by the vAPP layer	=3Mi CPU, =60MB RAM	Both UCs	Internal Metrics	Absolute Evaluation	RAM size, CPU cycles required

Table 8. SAFE-6G Safety function Core KPIs.

KPI category	*KPI name*	*KPI definition*	*KPI target value*	*Use case*	*Data sources*	*KPI evaluation*	*Verification method/tools*
Latency	Tunnel latency	E2E delay between UE and target application. Measured on safety function level (excluding physical layers).	<15 ms	Both UCs	SDPGW	Absolute	Timestamp with Packet inspection on the GW
Latency	Tunnel Jitter	E2E fluctuation delay between UE and target application	<2 ms	Both UCs	SDPGW	Absolute	Timestamp with Packet inspection on the GW
Operational Network	Tunnel Packet Loss	E2E packet loss between UE and target application	<0.1%	Both UCs	SDPGW	Absolute	SNMP features on the GW
Operational Network	Tunnel establishment attempt	Number of requests for tunnel creation on SDPGW	<5 per UE/day	Both UCs	SDPGW	Absolute	Counting the number of attempts on GW
Operational Network	Tunnel establishment success	Percentage of successful tunnel establishments on SDPGW	>90%	Both UCs	SDPGW	Relative	Division of successful /total number of attempts on GW
Operational Network	Tunnel establishment failure	Percentage of failures tunnel establishments on SDPGW	<10%	Both UCs	SDPGW	Relative	Division of failed/total number of attempts on GW

Table 9. SAFE-6G Safety function Extended KPIs.

4.1.2.4 SAFE-6G RESILIENCE FUNCTION KPIS

KPI category	*KPI name*	*KPI definition*	*KPI target value*	*Use case*	*Data sources*	*KPI evaluation*	*Verification method/tools*
Latency	Resilience Score Calculation Latency	Measure the time taken from data input (service and CNC) to the calculation of the updated Resilience Score.	≤ 1 second	Ensuring timely updates to the Resilience Score.	Core logs, performance monitoring tools.	**Absolute** evaluation of calculation latency.	Performance testing and log analysis.
Latency	Resilience Action Recommendation Latency	Measure the time taken to process data from the moment it is received to the point where a resilience recommendation is made.	≤ 3 second	Ensuring timely resilience recommendations in near real-time scenarios.	Network traffic logs, DSS processing logs.	**Absolute**	Performance testing and log analysis during peak traffic periods.
Reliability and Availability	Resilience Function Uptime	Measure the percentage of time the Resilience Function is operational and available.	≥ 99%	Ensuring high availability of the DSS for continuous operation.	System uptime logs, monitoring tools.	**Relative**	Monitoring system uptime logs over a specified period.

Table 10. SAFE-6G Resilience function Core KPIs

KPI category	*KPI name*	*KPI definition*	*KPI target value*	*Use case*	*Data sources*	*KPI evaluation*	*Verification method/tools*
Performance	AI Orchestrator Satisfaction Rate	Measure the satisfaction of the desired Resilience Scores requested by the AI Orchestrator.	≥ 80% satisfaction	Ensuring the Resilience Score suggestions meet orchestrator expectations.	Function Logs.	**Absolute** evaluation of system uptime.	Conducting periodic manual audits and log analysis.
Resilience Trust Function	Level of Resilience	The time required for the Level of Resilience to be provided by the TF to the Cognitive Coordinator.	<= 3 min	Both UCs	Internal Metrics	Absolute Evaluation	Time elapsed

Table 11. SAFE-6G Resilience function Extended KPIs

4.1.2.5 SAFE-6G RELIABILITY FUNCTION KPIS

KPI category	*KPI name*	*KPI definition*	*KPI target value*	*Use case*	*Data sources*	*KPI evaluation*	*Verification method/tools*
Performance and Accuracy	Number of metrics	The number of metrics (by network and application layers) that can be used and processed by the Reliability function	≥ 10	Both UCs	Log files from Reliability function.	Absolute	Number of metrics count
Performance and Accuracy	Type of alerts	The different type of alerts that can be produced by the Reliability function	≥ 3	Both UCs	Log files from Reliability function	Absolute	Number of alerts count
NFV energy efficiency	NFV green deployment	Support the service deployment into an infrastructure powered by renewal energy sources to decrease the CO_2 emissions	15% reduction	Both UCs	Historic of power consumption metric from Kepler service, provided by the continuum.	Relative	Energy consumption of running service
ML energy efficiency	Energy aware ML models	Development of energy aware versions of ML models that can reduce energy consumption during training/inference	20% reduction	Both UCs	Historic of power consumption metric from Kepler service, provided by the continuum.	Relative	Energy consumption during training/inference
Reliability Trust Function	Level of Reliability	The time required for the Level of Reliability to be provided by the TF to the Cognitive Coordinator.	<= 3 min	Both UCs	Internal Metrics	Absolute Evaluation	Time elapsed

Table 12. SAFE-6G Reliability function Extended KPIs

4.1.3 SAFE-6G CLOUD CONTINUUM KPIS

KPI category	*KPI name*	*KPI definition*	*KPI target value*	*Use case*	*Data sources*	*KPI evaluation*	*Verification method/tools*
Operational Network	aerOS Service availability	It represents the ratio between the amount of time during which aerOS services are healthy and running, thus fulfilling the expected QoS requirements.	>99,9%	UC agnostic	Historic from K8s' relevant metrics stored in the DataOps (from kube-state-metrics, Gatus or similar agent)	Absolute	Mean of the percentage of time aerOS' services are healthy
Operational Network	aerOS Service reliability	It represents the ratio between the number of requests aerOS services are handling, with respect to all the petitions received.	>99,9%	UC agnostic	Observability agent metrics (Cilium, Hubble) stored in DataOps, with information about requests, success, and error rate	Absolute	100*(Number of responses / number of requests) of aerOS services
Capacity	aerOS Orchestration success rate	Percentage of successful requests handled to allocate and deploy services, from the API or the portal.	>90%	Both UCs	Observability agent metrics (Cilium, Hubble) stored in DataOps, monitoring requests and error rate of the HLO endpoint	Absolute	100*(Number of successful responses/ numbers of accepted requests) of HLO
Operational Network	Function creation success	Percentage of successful function creations through the meta-OS.	>90%	Both UCs	Logs of aerOS' HLO	Absolute	100*(Number of successful responses/ number of requests)
Operational Network	SAFE-6G Function Uptime	Measure the percentage of time the Trust Functions are	≥ 99.9% (all)	Ensuring high availability of the DSS.	Historic from K8s' relevant metrics stored in the DataOps	Absolute evaluation of system uptime.	Mean of the percentage of time trust functions are healthy

		operational and available.			(from kube-state-metrics, Gatus or similar agent)		

Table 13. SAFE-6G aerOS Core KPIs

KPI category	*KPI name*	*KPI definition*	*KPI target value*	*Use case*	*Data sources*	*KPI evaluation*	*Verification method/tools*
Energy	NFV green deployment	Support the service deployment into an infrastructure powered by renewal energy sources to decrease the CO_2emissions.	15% reduction	Both UCs	The historic power consumption from Kepler agent is stored in DataOps. Metadata available at data fabric for identifying energy source from domains' IEs	Relative	Monitoring of the evolution of such consumption, when computing element is changed
Latency	aerOS Orchestration time	Time required for the orchestration of IoT applications (from command to selection of IE).	<10 s	Both UCs	Observability agent metrics (Cilium, Hubble) stored in DataOps, which has information about response time	Absolute	Average of the gathered response time
Operational Network	Function creation attempt	Number of requests for function creation through the meta-OS.	<5 per UE/day	Both UCs	Logs of aerOS' HLO	Absolute	Number of API requests to HLO API in specific time window

Table 14. SAFE-6G Extended KPIs

4.1.4 SAFE-6G CORE KPIS

KPI Category	*KPI name*	*KPI definition*	*KPI target value*	*Use Case*	*Data sources*	*KPI evaluation*	*Verification method/tools*
Other	UE positioning	Accuracy of the UE positioning. This is needed for REQ-USER-UC1-F-R-1 to get the positioning of users, machines and sensors within the factory.	Most accurate level: <= 5m for UC1 if possible. Cell-ID if gNB does not support NRPP	UC1	NA	Absolute	Compare the 5G positioning with GPS reference values (applicable outdoors) Check the Cell-ID is reported when UE changes gNB
Operational Network	Use Cases Reliability	Packet loss rate / frame loss / System uptime	UC1: comparable to URLLC systems (Desired<< 0.1%)	Both UCs	NA	Relative	Measure packet delivery with iperf3 application in UE and server
Latency	End-to-end latency	End-to-end latency between 2 UEs connected to the same Safe-6G network and involved in a Use-case	< =15ms required for UC1, <=20ms desired for UC2	Both UCs	End-to-end latency -> measure of time for a data payload sends by UE1 (ex: XR headset) to be received by UE2 (another headset). Reference points need to be discussed depending on how XR headsets are connected to the 6G network (ex: intermediate Wi-Fi hotspot...)	Absolute	Measure with ICMP the RTT delay and with iperf3 delay jitter to check delay variation
Operational Network	Use Case Availability	Service availability	>= 99.99% for Both use cases	Both UCs	See above	Absolute	Check the system uptime

Latency	Latency Jitter	Latency jittering is far more detrimental for real-time interactions in XR => better to have a higher but stable latency than having spikes.	Lowest achievable (is +- 5-10ms <= 5ms for UC1, <= 10ms for UC2	Both UCs	Same as latency, see above	Absolute	Measure with ICMP the RTT delay and with iper3 the delay jitter
Capacity	End-to-end bandwidth	It depends on which data is used and which technology. It depends on whether services like Nvidia CloudXR are used or not. UC1: potentially 2-3 HD video streams, real-time data from IOT, animated 3D models in XR. UC2: For each user: (3-5 users): 1 HD video+audio stream, XR avatars animated in real time.	>= 100Mbps downlink and 50 Mbps uplink	Both UCs	See above	Absolute	Measure with iperf3 the throughput using TCP and UDP stream sessions with iperf3 client in UE and iperf3 in server connected to the 5GC

Table 15. SAFE-6G General KPIs.

4.1.5 SAFE-6G MLOPS FRAMEWORK KPIS

KPI Category	*KPI Name*	*KPI Definition*	*KPI target value*	*Use Case*	*Data Sources*	*KPI Evaluation*	*Verification Method/Tools*
Deployment	Deployment Time of MLOps Framework	The total time taken to deploy the MLOps framework, from initialization to full operational readiness.	<15min	Both UCs	Data Sources defined by TFs	Absolute	Measure the elapsed time using timestamps recorded at the start of deployment and completion, ensuring functionality.
Scalability	Scalability Setup Time	Time taken to expand infrastructure during high-demand scenarios.	<15min	Both UCs	Data Sources defined by TFs	Absolute	Measure the time between scaling initiation and operational readiness of resources, validated

							through monitoring tools.
Availability	System Uptime	Percentage of time the MLOps infrastructure is operational and functioning as expected without interruptions.	>80%	Both UCs	Data Sources defined by TFs	Relative	Monitor the system using uptime monitoring tools, calculating uptime percentage over a period.
Compatibility	AI/ML Frameworks Supported	The number and variety of AI/ML frameworks supported by the MLOps infrastructure.	At least 3	Both UCs	Data Sources defined by TFs	Absolute	List supported frameworks, validate integration with the platform's pipeline, ensuring compatibility for deployment.
Security	Multi-Tenant Isolation Compliance	Measures the level of isolation between different users within the MLOps platform.	99%	Both UCs	Data Sources defined by TFs	Relative	Review access control configurations, conduct tests, monitor audit logs, and perform penetration testing.
Automation	Automation Rate for MLOps Infrastructure	Measures the extent to which the MLOps infrastructure is designed to automate key aspects of the machine learning lifecycle.	>60%	Both UCs	Data Sources defined by TFs	Relative	Calculate automation percentage using the formula: (automated processes / total processes) * 100.

Table 16. SAFE-6G AI and MLOps KPIs

4.1.6 SAFE-6G XAI KPIs IN MLOPS

KPI category	*KPI name*	*KPI definition*	*KPI target value*	*Use case*	*Data sources*	*KPI evaluation*	*Verification method*
Explainability	Number of local XAI methods	Number of local XAI methods to be developed for the project to answer the diverse needs of explainability	>=3	Both UCs	dataset + new data, AI/ML model	Absolute	Absolute
Explainability	XAI Method selection precision	Precision and pertinence of the selected XAI methods by the XAI assistant during interactions with the end-users	Should be refined during the project life.	Both UCs	dataset + new data, AI/ML model	Absolute	Expert

		This KPI has also been proposed as a project KPI					

Table 17. SAFE-6G XAI KPIs in MLOPs

4.1.7 SAFE-6G Differential Privacy KPIs in MLOPS

KPI category	*KPI name*	*KPI definition*	*KPI target value*	*Use case*	*Data sources*	*KPI evaluation*	*Verification method*
Differential Privacy	Accuracy Drop due to Differential Privacy	Measures the decrease in model accuracy when DP is applied compared to model accuracy with no DP.	≤ 10% accuracy drop	Both UCs	Model evaluation metrics before and after applying DP.	Compare accuracy of models trained with and without DP, using standard ML evaluation tools.	Relative comparison

Table 18. SAFE-6G differential Privacy KPIs in MLOps

4.1.8 SAFE-6G Chatbot KPIs

KPI category	*KPI name*	*KPI definition*	*KPI target value*	*Use case*	*Data sources*	*KPI evaluation*	*Verification method*
Chatbot	Accuracy of Intent Recognition	Measures the chatbot's ability to correctly understand and classify user intents. Evaluate a sample of interactions where the chatbot's intent recognition is compared against human annotations. Calculate the percentage of correctly recognized intents.	>80%	Both Ucs	Interaction logs	Relative evaluation	Use a small, fixed sample size for manual comparison
Chatbot	Fallback Rate	Tracks how often the chatbot resorts to a default response, indicating the need for additional training or scripting. Track the number of fallback responses used compared to the total number of interactions. Calculate the fallback rate as a percentage.	<20%	Both Ucs	Interaction logs	Relative evaluation	Use automated scripts to calculate fallback percentages.
Chatbot	AI Model Responsiveness	Monitors the speed at which the chatbot provides responses, affecting user satisfaction and efficiency.	<10sec	Both Ucs	System performance logs	Absolute evaluation	Set up automated response time

		Measure the time between a user query and the chatbot's response.					monitoring tools.
Chatbot	AI Model Error Rate	Measures the frequency of incorrect or irrelevant responses by the chatbot. Review a sample of chatbot interactions to identify incorrect or irrelevant responses. Calculate the error rate as a percentage of total interactions.	<5%	Both Ucs	Error logs	Relative evaluation	Randomly review flagged interactions monthly.
Chatbot	Task Completion Rate	Measures the rate at which users successfully complete tasks with the help of the chatbot. Track the number of tasks initiated and successfully completed with the chatbot's assistance. Calculate the completion rate as a percentage.	>80%	Both Ucs	User interaction logs	Relative evaluation	Use predefined scripts to track task completion rates.
Chatbot	Uptime	Measures the percentage of time that the chatbot is operational and available to users. Monitor the chatbot's operational status over a defined period and calculate the percentage of time it is up and running.	>99%	Both Ucs	Performance monitoring	Relative evaluation	Use automated uptime monitoring software with alerts.

Table 19. SAFE-6G Chatbot Core KPIs

KPI category	*KPI name*	*KPI definition*	*KPI target value*	*Use case*	*Data sources*	*KPI evaluation*	*Verification method*
Chatbot	User satisfaction	Assessed through surveys or feedback mechanisms to gauge users' feelings about their interactions with the chatbot.	>85%	Both Ucs	Survey results	Relative evaluation	Regularly analyze survey feedback trends.
Chatbot	User Retention Rate	User Retention Rate measures how often users return to interact with the chatbot post-initial use	>65%	Both Ucs	User database	Relative evaluation	Check percentage of repeat users monthly

Table 20. SAFE-6G Chatbot Extended KPIs

5 SAFE-6G KVs & KVIs

The SAFE-6G framework introduces an innovative approach to evaluating and ensuring the alignment of 6G technologies with societal, environmental, and economic priorities through the adoption of **KVIs**. This approach prioritizes the four foundational KVIs defined before. These indicators serve to assess how effectively 6G innovations contribute to addressing societal challenges and advancing values that align with global objectives.

A key feature of the SAFE-6G framework is the integration of **Trustworthiness** metric, which is used to evaluate the dependability, security, and ethical alignment of the system. LoTw provides a structured mechanism for assessing how trustworthy a system is from multiple dimensions. This ensures that innovations within the 6G ecosystem meet both technical and ethical standards.

By combining these KVIs with the previous approach and a robust validation framework, SAFE-6G ensures a holistic assessment that not only advances technological capabilities but also reinforces the societal, environmental, and ethical foundations critical for the 6G era.

5.1 KV: TRUST

This KV emphasizes the identification and quantification of the impact that trust-related factors, such as safety, security, privacy, reliability, and resilience, have on system and user interactions. By determining the enablers and barriers that influence trust, this KV demonstrates how these elements support and enhance the value of trust and trustworthiness across various domains. Furthermore, it highlights the alignment of trust-related initiatives with global challenges outlined in the UN SDGs, such as fostering resilient infrastructure (SDG 9), ensuring inclusive, safe, and sustainable environments (SDG 11).

Through its comprehensive approach, the **KV: Trust** provides an approach to strengthen trust in technologies, processes, and systems while addressing critical societal and sustainability challenges.

SAFE-6G consortium aims to differentiate the terms “Trust” and “Trustworthiness” in 6G systems and explicitly define them, avoiding the often-observed interchangeable use of these terms by the community.

Trust is an attitude that a tenant has towards a 6G system. In contrast, **Trustworthiness** is a system property that creates trust to the 6G tenant/user. The more trustworthy the 6G system is, the higher the trust level of the tenant/user will be.

Template field	***Description***
KV Name	Trust
Use case	Ensuring secure, transparent, and reliable communication systems for users in Industry 4.0 environment and beyond. Ensuring secure, transparent, and reliable communication systems for users.
Stakeholders	End-users, businesses, regulators, and consumer advocacy groups.

Societal pain points	Data breaches, misinformation, lack of transparency, and low confidence in communication system reliability.
Positively impacted Key Values	Trust in communication systems.
Scale of effect: KVIs	Increased user confidence in the communication platform and assured access availability aligned with roles. Enhanced adoption rates due to perceived security and transparency.
Enablers and blockers	Enablers: Aligned with the SAFE6G Architecture, with a particular emphasis on the Level of Trustworthiness (LoTw), enabling dynamic adaptation of trustworthiness levels within the infrastructure. This approach ensures comprehensive implementation, addressing key dimensions such as resilience, security, privacy, reliability, and safety.
Quantification with KPIs	✓ LoTw

Table 21. Trust - SAFE 6G KVI reference description

5.1.1 KVI: TRUSTWORTHINESS

Trustworthiness is a system property that creates trust to the 6G tenant/user. To quantify this, the term **LoTw** is used, which represents the trustworthiness provided by the system to the tenant/user. The most significant paradigm adjustments in the envisioned user-centric 6G system are the shift from a security-only focus to a broader scope of native trustworthiness, clarifying that the term "trustworthiness" refers to a holistic approach, including safety, security, privacy, resilience and reliability as trustworthiness dimensions and properties. These five dimensions are implemented through the Trust Functions of the SAFE-6G framework. Each Trust Function, dedicated to a specific trustworthiness dimension, can receive inputs, make intelligent decisions, and, in collaboration with the Cognitive Coordinator, execute actions across three different planes:

- Application plane
- 6G Core network
- Resource edge/cloud continuum plane

This process ensures that the system transitions into a trustworthy state.

Each Trust Function provides a unique Level of TFx, which reflects the user's intent and is influenced by the system's current capabilities. An intelligent combination of the five distinct Levels of TFx, provided by the Cognitive Coordinator, defines the holistic Level of Trustworthiness, which is responsible for accommodating the user's intent, enhancing the user's trust in the system and realizes a user-centric trustworthy service provision over 6G, formulated as follows:

Formula:

$$LoTw = \sum_{j=1}^{N} w_{LoTF_j} \cdot LoTF_j$$

With the constraint:

$$\sum_{j=1}^{N} w_{LoTF_j} = 1.$$

Parameters LoTFj:

- Level of Security.
- Level of Privacy
- Level of Safety
- Level of Resilience
- Level of Reliability

Explanation:

- The Σ notation now sums over N Trust Functions.
- WLoTFj: represents the weight assigned to each Trust Function F Tj.
  The condition ensures that all weights sum to 1.
  Finally, it is stated that N = 5, making it clear that the summation is over five available Trust Functions provided by SAFE-6G.

LoTFj calculation:

- The LoTFj provided by each Trust Function towards the system, and the user, is calculated based on a rule based or intelligent method that maps specific inputs to specific flavors of actions. The value of the upper bound of the provided flavor defines the LoTFj.

From an implementation perspective, the computation of the Level of Trustworthiness (LoTw) is directly grounded on the KPI framework defined for the SAFE-6G Trust Functions in Section 3.2.2. Each Trust Function (Safety, Security, Privacy, Resilience, Reliability) continuously monitors technical indicators such as latency, uptime, accuracy, resource utilisation and security/privacy performance, using the system metrics and measurements that are also employed to derive its core and extended KPIs. These metrics constitute the input space of the rule-based or AI-driven methods that determine the per-function Level of TFx (LoTFj), which are then combined into the overall LoTw through the weighted aggregation defined above. In this way, LoTw acts as a KVI-level metric whose value is supported by concrete, measurable KPIs, and it is used as the quantification mechanism for KVIs such as “Trustworthiness” and “Privacy and Confidentiality” under the KVs “Trust” and “Security, privacy and confidentiality” (Table 16 - Table 17). This establishes an explicit traceability chain **from low-level KPIs → Trust Function levels (LoTFj) → system-level LoTw → KVs/KVIs**, ensuring that value-oriented indicators are systematically anchored to observable performance metrics.

5.2 KV: SECURITY, PRIVACY AND CONFIDENTIALITY

KV Security, Privacy and Confidentiality evaluates contributions to safeguarding sensitive information by identifying impacts, enablers, and barriers. It emphasizes how privacy and confidentiality support trustworthiness while aligning with UN SDGs. This KVI highlights the critical role of data protection in building trust and ensuring secure, sustainable systems.

Template field	*Description*
KV Name	Security, privacy and confidentiality
Use case	Protecting user data and ensuring confidentiality in communication systems.
Stakeholders	End-users, businesses, regulators, IT security teams, and data protection authorities.
Societal pain points	Unauthorized data access, lack of transparency in data handling, and breaches of confidentiality.
Positively impacted Key Values	Privacy and confidentiality of user communication and data.
Scale of effect: KVIs	Improved user confidence in data handling processes. Increased adherence to privacy regulations (e.g., GDPR). Fewer incidents of data breaches or unauthorized access.
Enablers and blockers	Enablers: End-to-end encryption, robust data protection policies, and regular audits. Blockers: Cyberattacks, inadequate employee training on data handling, and lack of compliance with privacy standards.
Quantification with KPIs	LoTw

Table 22. Security, privacy and confidentiality - SAFE 6G KVI reference description

5.2.1 KVI: PRIVACY AND CONFIDENTIALITY

It is inherently challenging to disentangle the concepts of trust, security, and privacy within a safety framework, as they are deeply interwoven and mutually reinforcing. Given this interdependence, the evaluation of these KVIs has been integrated into a unified metric known as the LoTw, which has been detailed in the previous section.

In this context, the KVI associated with Privacy and Confidentially aligns closely with the principles underlying the LoTw. This is an equivalent scenario like in Trustworthiness KVI. Since LoTw encapsulates essential aspects of security, privacy, and trust, the assessment criteria and considerations established in the earlier discussion remain applicable to this KVI. By leveraging the LoTw framework, a comprehensive and consistent approach is ensured for evaluating trust-related dimensions across different domains, including Environmental Sustainability.

5.3 KV: Sustainable Resource Utilization

For the Key Value Sustainable Resource Utilization, the KVI named *Environmental Sustainability* will be considered and elaborated in the following subsection.

The ***KVI: Environmental Sustainability*** measures contributions to sustainable practices by evaluating impacts, enablers, and barriers. It emphasizes resource optimization, energy efficiency, and resilience while aligning with UN SDGs, such as climate action (SDG 13) and sustainable cities (SDG 11). This KVI highlights the role of environmentally responsible systems in promoting sustainability and long-term resilience.

Template field	*Description*
KV Name	Environmental Sustainability
Use case	Adoption of green energy sources and waste management systems.
Stakeholders	Environmental agencies, supply chain managers, and consumers.
Societal pain points	High carbon emissions, unsustainable resource consumption, and waste overflow.
Positively impacted Key Values	Environmental sustainability.
Scale of effect: KVIs	Lower carbon footprint and reduce waste. Improved compliance with environmental regulations.
Enablers and blockers	Enablers: Renewable energy sources, recycling technology. Blockers: Excessive costs of green technologies, regulatory complexities.
Quantification with KPIs	✓ Resource Utilization Reduction Rate >10% ✓ 6GEnergyEfficiencyImprovement >20% ✓ Resilience Actions Successful Rate >90%

Table 23. Sustainable Resource Utilization - SAFE 6G KVI reference description

5.3.1 KVI: Sustainable Resource Utilization

The Relative Utilization of Resource Reduction (RURR) is an important metric designed to evaluate the efficiency of resource utilization within a system. It specifically measures the difference between the maximum energy capacity the system is designed to handle (*SystemMaxEnergyUsed*) and the actual average energy consumed during operations (*SystemAvgRealEnergyUsed*), expressed as a percentage of the maximum capacity. By quantifying this difference, the RURR provides insights into the system's operational efficiency, highlighting scenarios of energy savings or overconsumption. This metric not only captures the baseline resource usage but also accounts for variations in energy performance, offering a robust and scalable framework to monitor and optimize resource utilization over time.

Parameters:

- *SystemMaxEnergyUsed*: Represents the maximum energy the system is designed to utilize during operation.

- *SystemAvgRealEnergyUsed*: Denotes the actual average energy consumption of the system during operation.

Formula:

$$\text{RURR}\ (\%) = \frac{\text{SystemMaxEnergyUsed} - \text{SystemAvgRealEnergyUsed}}{\text{SystemMaxEnergyUsed}} \times 100$$

- If *SystemAvgRealEnergyUsed* is less than *SystemMaxEnergyUsed*, the RURR value will be a positive percentage, indicating a reduction in energy usage compared to the system's maximum design capacity. And conversely, if *SystemAvgRealEnergyUsed* is greater than *SystemMaxEnergyUsed*, the RURR value will be a negative percentage, signifying an increase in energy consumption beyond the system's intended maximum capacity.

5.3.2 KVI: Resource Provisioning Efficiency

For the Key Value Environmental Sustainability, the KVI named *Resource Provisioning Efficiency* will be considered and elaborated in the following subsection.

The *Resource Provisioning Efficiency (RPE)* is a key metric that refers to the optimization of resources allocated to perform a specific task or group of tasks. It averages the KPIs *CPUEfficiencyRatio* and *RAMEfficiencyRatio*, which represent the average ratio of resources in use with (CPU_{actual}, RAM_{actual}) respect to those allocated (CPU_{prov}, RAM_{prov}) to support the performance of the deployed services. While there are no official values, a consensus is to have a ratio of CPU usage between 60-80% and 20-70% of RAM (since it is a cheaper resource). By providing insights about resource provisioning, RPE helps optimize the configuration of services and automatic allocation rules.

Parameters:

- *CPU in use (**$CPU_{act,kn}$**)*: CPU consumed by service K in time N.
- *RAM in use (**$RAM_{act,kn}$**)*: RAM consumed by service K in time N.
- *CPU provisioned (**$CPU_{prov,kn}$**)*: CPU provisioned to service K in time N.
- *RAM provisioned (**$RAM_{prov,kn}$**)*: RAM provisioned to service K in time N.

$$CPUEfficiencyRatio(\%) = \frac{100}{K \cdot N} \sum_{k=1}^{K} \sum_{n=1}^{N} CPU_{kn}$$

$$CPU_{kn} = \begin{cases} 1.0, & 0.6 \leq \frac{CPU_{act,kn}}{CPU_{prov,kn}} < 0.8 \quad \text{(Optimal)} \\ 0.5, & 0.3 \leq \frac{CPU_{act,kn}}{CPU_{prov,kn}} < 0.6 \text{ or } 0.8 \leq \frac{CPU_{act,kn}}{CPU_{prov,kn}} < 0.9 \quad \text{(Suboptimal)} \\ 0.0, & \text{otherwise} \quad \text{(Poor)} \end{cases}$$

$$RAMEfficiencyRatio(\%) = \frac{100}{K \cdot N} \sum_{k=1}^{K} \sum_{n=1}^{N} RAM_{kn}$$

$$RAM_{kn} = \begin{cases} 1.0, & 0.2 \leq \frac{RAM_{act,kn}}{RAM_{prov,kn}} < 0.7 \quad \text{(Optimal)} \\ 0.5, & 0.1 \leq \frac{RAM_{act,kn}}{RAM_{prov,kn}} < 0.2 \text{ or } 0.7 \leq \frac{RAM_{act,kn}}{RAM_{prov,kn}} < 0.8 \quad \text{(Suboptimal)} \\ 0.0, & \text{otherwise} \quad \text{(Poor)} \end{cases}$$

$$RPE(\%) = \frac{CPUEfficiencyRatio + RAMEfficiencyRatio}{2}$$

The closer the RPE value is to 100%, the better allocation of resources has been made. A value lower than 50% would imply the need for better configuring the resources allocated to serve the service, either because of having overprovisioned or under provisioned.

5.3.3 KVI: Sustainable Cities and Communities

For Key Value Energy Efficiency, the KVI named *Sustainable Cities and Communities* will be considered and elaborated in the following section.

The **Resilience Function Score (RFS)** is a key metric that evaluates the effectiveness of a system's resilience actions. It calculates the percentage ratio between successfully executed resilience actions and the total attempts made, providing a clear measure of the system's ability to respond effectively to resilience demands. By highlighting the success rate of resilience actions, RFS offers valuable insights into system performance and supports efforts to enhance resilience strategies over time.

$$RASR = \text{RFS} = \frac{\text{Resilience Actions Success}}{\text{Resilience Actions Attempts}} \times 100$$

Parameters:

1. *Resilience Actions Success*: The total number of successfully executed resilience actions.
2. *Resilience Actions Attempts*: The total number of attempted resilience actions.
3. *RFS*: The success rate of resilience actions, expressed as a percentage (0 to 100), calculated as the ratio of successful actions to total attempts.

Explanation:

1. *RFS*: This metric reflects the effectiveness of resilience actions by comparing the number of successful actions to the total number of attempts. For instance, if there are 90 successful actions out of 100 attempts, the RFS is *90%*.
2. Purpose: The RFS provides a more precise assessment by incorporating both successful outcomes and total attempts, ensuring a realistic evaluation of resilience performance.

5.4 KV: Simplified Life

The *KVI: Simplified Life* evaluates contributions to creating user-friendly, efficient, and accessible systems by assessing impacts, enablers, and barriers. It focuses on reducing complexity, improving

usability, and enhancing accessibility, aligning with UN SDGs, such as quality education (SDG 4) and reduced inequalities (SDG 10). This KVI highlights the importance of simplifying interactions to promote inclusion and improve quality of life.

Template field	*Description*
KV Name	Simplified Life
Use case	Development of intuitive, user-friendly systems and services that reduce complexity, enhance usability, and improve accessibility for all users.
Stakeholders	End-users, businesses, service providers, policymakers, and accessibility advocacy groups.
Societal pain points	• High complexity in technology and services. • Limited accessibility for people with disabilities or low technical literacy. • Time-consuming and inefficient processes in daily life.
Positively impacted Key Values	Enhanced accessibility and inclusion. Improved user satisfaction and quality of life.
Scale of effect: KVIs	• Increased adoption of simplified systems due to reduced complexity. • Greater accessibility, ensuring inclusivity for diverse populations. • Enhanced productivity and efficiency in personal and professional contexts.
Enablers and blockers	Enablers: Advances in human-centered design, AI-driven personalization, and simplified interfaces. Blockers: Resistance to change, lack of investment in accessibility features, and technological disparities across regions.
Quantification with KPIs	✓ NetworkSafeActionsAutoRatio > 30%

Table 24. Simplified Life - SAFE 6G KVI reference description

5.4.1 KVI: Dynamic Network Programmability

For the Key Value Simplified Life, the KVI named *Dynamic Network Programmability* will be considered and elaborated in the following section.

The **Network Safe Actions Auto Ratio (NSAAR)** is a key metric designed to measure the proportion of automated actions within a system's overall operations. It evaluates the relationship between the total number of automatic actions (*TotalAutoActions*) and the total actions performed (*TotalActions*), including both automated and user-initiated actions. By quantifying this ratio, NSAAR provides insights into the system's reliance on automation, highlighting the balance between efficiency through automation and manual intervention. This metric plays a vital role in assessing operational performance, ensuring a safe and effective distribution of system processes while supporting efforts to enhance automation where appropriate.

Parameters:

- *SystemAutoActionsCount*: Total number of automatic actions performed by the system.
- *SystemAutoActionsSuccess*: Count of successful automatic actions.
- *SystemAutoActionsFail*: Count of failed automatic actions.
- *SystemRequestedActionsCount*: Total number of requested (manual or user-initiated) actions.
- *SystemRequestedActionsSuccess*: Count of successful requested actions.
- *SystemRequestedActionsFail*: Count of failed requested actions.

Formulas:

- Total Auto Actions:
  TotalAutoActions = SystemAutoActionsCount
- Total Requested Actions:
  TotalRequestedActions = SystemRequestedActionsCount
- Total Actions:
  TotalActions = TotalAutoActions + TotalRequestedActions
- Network Safe Actions Auto Ratio (NSAAR):
  NSAAR = TotalAutoActionsTotalActions / TotalActions

Explanation:

- Maximum Energy Used (E_{max}): The highest possible energy consumption under full load or operational capacity.
- Average Real Energy Used (E_{avg}): The actual energy consumption averaged over a specific period or usage cycle.

5.5 KV: DIGITAL INCLUSION

The KVI: Digital Inclusion evaluates efforts to bridge the digital divide by ensuring equitable access to digital technologies and services for all, regardless of geographic, socioeconomic, or physical barriers. It emphasizes the development of scalable, affordable, and accessible 6G networks, aligning with UN SDGs, such as reduced inequalities (SDG 10) and quality education (SDG 4). This KVI highlights the role of inclusive technologies in empowering underserved communities, fostering digital literacy, and enabling participation in digital economies, contributing to a more equitable and connected society.

Template field	***Description***
KVI Name	Digital Inclusion
Use case	Development of affordable, scalable, and universally accessible 6G services and infrastructure to bridge the digital divide and promote inclusivity across diverse populations.
Stakeholders	End-users, rural and underserved communities, policymakers, service providers, and technology developers.

Societal pain points	• Limited access to digital technologies and services in rural or underprivileged areas. • Excessive costs of digital infrastructure and services, creating barriers for low-income groups. • Lack of digital literacy hinders the effective use of available resources.
Positively impacted Key Values	Enhanced accessibility and inclusion. Improved user satisfaction and quality of life.
Scale of effect: KVIs	• Increased adoption of simplified systems due to reduced complexity. • Greater accessibility, ensuring inclusivity for diverse populations. • Enhanced productivity and efficiency in personal and professional contexts.
Enablers and blockers	Enablers: Advances in human-centered design, AI-driven personalization, and simplified interfaces. Blockers: Resistance to change, lack of investment in accessibility features, and technological disparities across regions.
Quantification with KPIs	✓ The quantification of this KVI is aligned with societal inclusion and must be defined according to the real scenario of the deployment.

Table 25. Digital Inclusion - SAFE 6G KVI reference description

5.6 KV: SOCIETAL SUSTAINABILITY

The *KV: Societal Sustainability and Innovation* measures contributions to fostering sustainable growth and driving innovation by evaluating impacts, enablers, and barriers. It focuses on optimizing resources, promoting efficiency, and encouraging innovative practices, aligning with UN SDGs, such as decent work and economic growth (SDG 8) and industry, innovation, and infrastructure (SDG 9). This KVI underscores the importance of sustainable economic strategies to ensure long-term prosperity and resilience.

Template field	***Description***
KVI Name	Data protection Index
Use case	Ensuring secure and equitable access to systems that protect privacy, confidentiality, and transparency to support democratic processes and values.
Stakeholders	Citizens, policymakers, governments, non-governmental organizations (NGOs).
Societal pain points	• Lack of privacy and confidentiality undermining democratic participation. • Inequitable access to secure systems and information. • Erosion of trust in democratic institutions due to data breaches or misinformation.
Positively impacted Key Values	• Strengthened trust in democratic systems and processes. • Enhanced transparency and accountability in governance.

	• Empowered citizen participation.
Scale of effect: KVIs	• Increased trust in governance and democratic processes through secure systems. • Improved access to digital platforms for equitable participation in democracy. • Strengthened protections for privacy and freedom of expression.
Enablers and blockers	Enablers: Advanced data protection technologies, transparency frameworks, and adherence to privacy and confidentiality standards. Blockers: Misinformation campaigns, lack of regulatory enforcement, and digital inequality among populations.
Quantification with KPIs	Data Protection Index >90%

Table 26. Societal Sustainability and Innovation - SAFE 6G KVI reference description

5.6.1 KVI: DATA PROTECTION INDEX

The **Data Protection Index (DPI)** is a key metric that measures the overall effectiveness of a system's data protection capabilities. It incorporates inputs such as the proportion of encrypted users (*EE*E) and the impact of security incidents (*SS*S and *S100*), accounting for their respective weights and influence. By balancing these factors, the DPI provides a comprehensive evaluation of data security performance. A higher DPI reflects robust data protection, with strong encryption practices and minimal security incidents, while a lower DPI indicates potential vulnerabilities or areas needing improvement. This metric enables organizations to monitor, assess, and enhance their data protection strategies over time.

Formula:

$$\mathrm{DPI} = w_E \times EEE - \left(w_S \times \frac{SSS}{S_{\max}} + w_{S100} \times \frac{S100}{S100_{\max}} \right)$$

Parameters:

- *Encrypted Users Ratio (EEE):*
  - Definition: The proportion of users whose data is encrypted.
  - Impact on DPI: Positive (higher EEE = stronger data protection).
  - Weight: wE (adjustable to reflect encryption's relative importance).
- *Security Incident Count (SSS):*
  - Definition: The total number of security incidents detected.
  - Impact on DPI: Negative (more incidents = weaker data protection).
  - Weight: wS (determines the significance of incident frequency).

- *Security Incidents per 100 Users (S100):*
    - Definition: The number of security incidents normalized per 100 users.
    - Impact on DPI: Negative (higher rate = poorer performance).
    - Weight: wS10 (reflects the impact of incident density).

Explanation:

- **Positive and Negative Contributions:** The positive effect of encryption coverage (EEE) is offset by the negative impact of security incidents (SSS and S100), ensuring a balanced performance assessment.
- **Normalization:** Security incident metrics are normalized to a 0–1 scale, enabling comparability across systems of varied sizes and environments.
- **Weighted Impact**: The weights (wE, wS, wS100) are customizable, allowing organizations to prioritize certain factors depending on their security strategies or operational contexts.

6 VALIDATION METHODOLOGY

6.1 DEVELOPMENT, INTEGRATION, AND VERIFICATION METHODOLOGY

SAFE-6G verification aims to guarantee that the components delivered by the technical work packages during partial system integration phases meet the set of design specifications defined during the architectural specification phase and cover the respective items of the final platform. During intermediate development and partial system integration phases, verification procedures will be developed. Such verification procedures should define specific test scenarios modelling or simulating system components up to the complete SAFE-6G platform and methods for evaluating and analyzing the modelling results to guarantee operations.

SAFE-6G adopts a DevOps approach for development, integration, testing and deployment. DevOps is a set of practices that automate the processes between software development and IT teams, in order that they can build, test, and release software faster and more reliably. The concept of DevOps is founded on building a culture of collaboration between teams that historically functioned in relative siloes. The promised benefits include increased trust, faster software releases, and the ability to solve critical issues quickly and better manage unplanned work.

In SAFE-6G, DevOps philosophy and the corresponding approach will be used and applied internally for the development and the operation (deployment and validation) of the software components in the project as SAFE-6G platform is composed of a set of software components that will be implemented by different partners following different technologies. To overcome the integration challenges, SAFE-6G will use a DevOps based approach to be able to fully support the management of these implementations and the planned releases. DevOps integrates development and operations into a single-minded entity with common goals: high-quality software, faster releases, and improved users' satisfaction. This approach also incorporates several agile principles, methods, and practices, such as Continuous Delivery, Continuous Integration, and Collaboration.

Continuous Integration is a developer practice to keep a working system through minor changes and grow the system by integrating frequently (usually at least daily) on the mainline by means of appropriate tools supporting automation with lots of automated tests. This enables teams to work on shared code and increases the visibility into the development and quality of the system.

SAFE-6G will be based on open-source projects and tools from Cloud Native Computing Foundation (CNCF), and the results of previous H2020 projects. The latest version of the software will be automatically integrated upon successful source code updates and compatibility tests, and an updated version will be deployed on the Integration Infrastructure hosted in the UNIWA, NCSRD, and UPV premises. Before each major or minor release cycle, the Staging Infrastructure will be used for extensive functional and penetration testing and bug fixing without interfering with the development of new releases happening in the Integration Infrastructure; in this way, at release time, the pilots will be updated with zero downtime.

For a successful implementation of a DevOps approach, automated development and delivery pipelines that consist of the stages an application goes through from development to production are required, as shown in the figure below. The Figure 9 shows the environments that are envisioned in SAFE-6G, covering the different development stages.

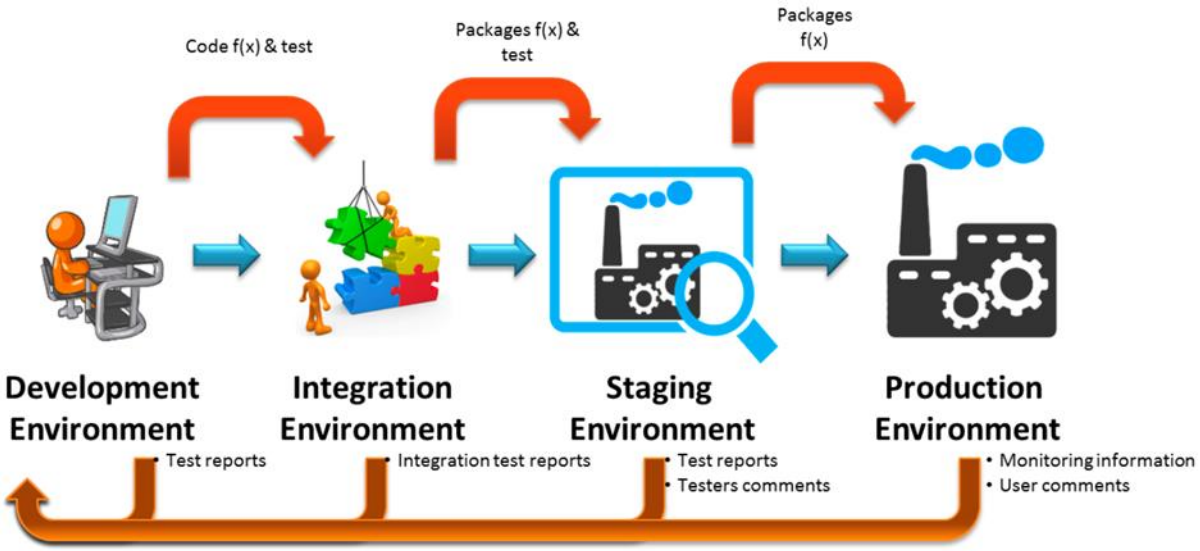

Figure 9. Environments envisioned in SAFE-6G.

6.2 Testing Levels

It is worth mentioning that component integration and verification/testing activities of the software components are executed before their deployment into the holistic scenarios. The following test levels will be defined:

Component testing (also known as unit, module, or program testing) searches for defects in and verifies the function of software modules, programs, objects, classes, etc., that are separately testable. It may be done in isolation from the rest of the system, depending on the context of the development life cycle and the system. Stubs, drivers, and simulators may be used. In the context of SAFE-6G platform development separate component tests will be planned and executed to facilitate the verification at component level (unit-test).

Integration testing validates interfaces between components, interactions with various parts of a system and interfaces between systems. Systematic integration strategies may be based on system architecture (such as top-down and bottom-up), functional tasks, transaction processing sequences, or some other aspect of the system or components. To ease fault isolation and detect defects early, integration should normally be incremental rather than “big bang”.

System testing is concerned with the behavior of a whole platform. In system testing, the test environment should correspond to the final target or production environment as much as possible to minimize the risk of environment-specific failures not being found in testing. System testing may include tests based on risks and/or on requirements specifications, business processes, use cases, or

other high-level text descriptions or models of system behavior, interactions with the operating system, and system resources. In SAFE-6G, this level of testing should be scheduled during product releases and is expected to be facilitated by the SAFE-6G demonstrators.

Other indicative test types/categories are listed below:

- **Functional testing.** It tests a slice of functionality in a system. The slice of functionality can be a unit of system behavior; a complex behavior composed of many units' micro-services and can be the whole system's behavior. It aims to test whether the expected behavior is successfully done by the system rather than how the behavior is done with which quality.
- **Performance testing.** It is a non-functional testing technique performed to determine the system parameters in terms of responsiveness and stability under various workloads. Performance testing measures the quality attributes of the system, such as scalability, reliability, latency and resource usage.
- **API testing.** It is part of the functional testing that aims to test whether the implemented API behaves as expected in the specifications.
- **System Testing.** It is a comprehensive system testing to evaluate the overall functionality of SUT. This type of testing verifies that the system meets all the specified requirements and performs as expected in terms of robustness and reliability.

The tests will follow a scenario-driven approach, and each scenario will cover a part of the integration workflow. The specified tests might be distinguished in bilateral, that is between components A and B, and/or system-level cross-cutting ones. The tests will be defined and report their outcomes according to a structured procedure. First, the scenario is defined. For this, the Scenario template presented below is used to specify the test. More specifically, the Scenario template incorporates details that pertain to the objective of the tests, the participating components, the requirements that are addressed, the features that are tested, the steps that need to be verified, and finally, the test setup, which provides details on the integration setting that facilitates the test. Then, the results are presented in detail as dictated by the steps that are defined in the scenario. Finally, the Checklist template is applied describing the successful or unsuccessful are reported and fall into the specific integration scenario.

<Test Name>
Objective
Involved Components
Requirements
Features to be tested
Test setup
Steps

Table 27. Testing scenario template

Test Checklist			
	Success	Fail	Comments
<test check point 1>			
<test check point 2>			
<test check point 3>			
<test check point 4>			

Table 28. Checkpoint templates

6.3 TEST PHASES

Each test procedure, performed at the staging environment of Figure 9 **¡Error! No se encuentra el origen de la referencia.**, consists of several steps that go through during the execution. The tests can be triggered manually by the users, or they can be part of an automated sequence of actions that can be triggered automatically in a periodic way after a specific action, i.e., push of new code, release of new component version, etc. All tests will be fully integrated with the Continuous Integration/Continuous Delivery (CI/CD) framework and implemented as executable scripts by well-known automation tools like GitLab runners, Jenkins servers, Ansible, etc. In its general form, a test will consist of the following phases:

- **Test Preparation**. In test preparation, the test environment must be set up; this involves deploying an instance of the under-test component in the test environment (i.e., stage, integration, etc.) and loading all the necessary additional test libraries and tools.
- **Test Execution**. Once the environment is configured and set up the tests are executed in a serial way to ensure that the generated results are not affected by parallel test executions.
- **Documentation and Reporting**. All test plans, test cases, test results, and any issues encountered during testing are documented. Comprehensive reports that provide an overview are generated and provided to relevant stakeholders.
- **Test Shutdown**. Once the tests are completed, all the allocated resources (i.e., containers, K8s pods, instances of testing and benchmarking tools, memory, etc.) are released, and the system is ready for the execution of the next test.

6.4 SAFE-6G PLATFORM VALIDATION

Automated CI/CD pipelines will play a significant role in the evaluation of the SAFE-6G KPIs by integrating monitoring and testing mechanisms into CI/CD workflows, which can provide real-time insights into platform performance, reliability, and usability.

The SAFE-6G methodology is based on the following key features:

- Usage of state-of-the-art CI/CD tools (i.e. GitLab, Jenkins, CircleCI, etc) to schedule and execute automated test scripts for KPI collection.
- Integration of Continuous testing tools for performance evaluations (i.e. Apache JMeter, iPerf, Gatling, Locust), security testing (i.e. Trivy, Starboard, Kube-bench, etc.), energy (i.e., Kepler),

scalability and stress testing. Information from the continuum's data fabric related to resource utilization will also be managed.
- Development scripts for collecting KPIs values and integrating them into the CI/CD pipelines.
- Test results collection and Data analysis. Leverage CI/CD tools to automatically push KPI metrics into a centralized database for storage and analysis.
- Generation of reports on KPI compliance, and test failures in pipeline notifications.

6.5 Alignment with FAIR principles

SAFE-6G will publish ML datasets and the public deliverables related to the validation of the system and use-case KPI/KVIs on the project site, Zenodo, and the EOSC platform to align with the FAIR principles. Making deliverables publicly available on widely recognized repositories like Zenodo and EOSC enhances their visibility and discoverability through persistent identifiers and rich metadata. Hosting on these platforms supports open access, standardized formats, and community-agreed metadata standards, which improve interoperability and facilitate reuse by other researchers and stakeholders. This practice also demonstrates adherence to the FAIR ethos of maximizing the impact and transparency of research outputs, which is encouraged by European policies and research funding frameworks to accelerate scientific progress and innovation.

6.6 Platform Release planning

SAFE-6G will follow an agile and incremental approach to provide the intermediate and final release of the SAFE-6G ecosystem. The development of the components and tools will use an agile and incremental Continuous Integration/ Continuous Deployment/ Continuous Production (CI/CD/CP) approach, based on SCRUM methodology, allowing concurrent research, design, development, integration, deployment, testing, validation and qualification throughout the whole project, gradually providing an increasingly refined set of features, ultimately delivering the measurable KPIs as defined in SAFE-6G objectives. The release timeline is:

- Edge-cloud continuum infrastructure, Integration tools (Gitlab, CI/CD), MLOps framework.
- Integration of all modules and components (Cognitive Coordinator, Trust Functions).
- Intermediate release of the integrated SAFE-6G ecosystem.
- Final release of the Cognitive Coordinator and Trust Functions.
- Completion of the SAFE-6G system level test and verification.
- Final release of the SAFE-6G ecosystem.

7 TESTING TOOLS

This section describes the set of tools used to validate the KPIs and KVIs defined in this document. Each indicator is associated with one or more concrete tools, which are integrated into the SAFE-6G platform. The selection follows the TMV WG tool set definition, so that the project remains aligned with recognised industry practices and established performance evaluation methods.
All tools feed their outputs into the SAFE-6G observability and data-fabric components. From there, the results can be deposited in EOSC-compliant services for long-term storage, discovery, and reuse in line with FAIR data principles.

7.1 TOOL SET OVERVIEW AND KPI/KVI MAPPING

To cover KPIs and KVIs at the network, service, and societal levels, SAFE-6G uses a mix of active measurements, passive monitoring, security and privacy frameworks, and AI/ML observability tools. Together, they support end-to-end monitoring and validation of latency, throughput, security, privacy, resilience, AI/ML behaviour, and sustainability.
Each tool below is linked to its relevant KPIs/KVIs, with a short explanation of what it measures and how it does so.

1. **Latency Measurement Tools:**
   - Tool: [*iPerf]*
   - Applicable KPIs: Network Latency, Tunnel Latency, Response Time.
   - Description: These tools will be deployed across the SAFE-6G network layer to measure real-time latency metrics. Results will be logged and analyzed in the TMV-compliant monitoring module, enabling the correlation of latency data with other performance indicators.

What is measured
iPerf is used to characterise packet-level and flow-level delay on end-to-end paths, including intra- and inter-domain tunnels and links between key SAFE-6G components (for example, user-plane functions, edge nodes, and core services).

How it is measured
iPerf clients and servers are placed at representative points in the SAFE-6G network and generate controlled TCP/UDP traffic.

- *For Network Latency and Tunnel Latency, iPerf sends streams with configured packet sizes and sending rates. One-way and round-trip delays are derived from packet timestamps and sequence numbers.*
- *For Response Time, the iPerf session timestamps are correlated with logs from the relevant services in the TMV-compliant monitoring module to obtain end-to-end application-level delay.*

Measurements are repeated over time and under different load conditions, so that averages, percentiles, and jitter can be extracted.

Integration in SAFE-6G and EOSC

iPerf agents could be orchestrated through the SAFE-6G observability framework. The resulting metrics are exported in standard formats (for example, time-series data, CSV) into the data fabric and then deposited in EOSC-compatible repositories. Metadata describing the test scenario, topology, and configuration are attached to support discovery and reuse via EOSC services.

Note: additional tools could be identified during the progress of the project.

2. **Throughput and Capacity Validation:**
   - Tool: [*NetFlow, Prometheus]*
   - Applicable KPIs: Data Rates, Capacity, Tunnel Throughput.
   - Description: Throughput measurements will be conducted using network monitoring tools, which capture data rates and capacity utilization in real-time. The tool will interface with the SAFE-6G orchestration layer, providing insights into capacity metrics across distinct functions.

What is measured

These tools provide link-level, flow-level, and service-level throughput and capacity information across SAFE-6G network segments and tunnels, including interfaces that are critical for vertical services.

How it is measured

Network devices are configured with NetFlow (or equivalent exporters) to produce flow records containing byte and packet counters, start/stop timestamps, and QoS markings. By aggregating these flows per interface, tunnel, or service, Data Rates and Tunnel Throughput are obtained.
Prometheus scrapes metrics from network and service components (such as interface counters, queue occupancy, and rate gauges). Capacity is calculated by comparing measured utilisation with configured bandwidth and resource quotas per function. Different aggregation windows (e.g. 1 s, 10 s, 1 min) are used to capture average rates, peaks, and sustained throughput.

Integration in SAFE-6G and EOSC

NetFlow collectors and Prometheus exporters could be integrated with the SAFE-6G orchestration and monitoring layers. Metrics are normalised and stored in the data fabric, then offered to EOSC as curated datasets that include metrics, labels, and topological context. These datasets enable cross-project analysis and reproducibility.
Additional throughput and capacity tools can be added later in the project

Note: additional tools could be identified during the progress of the project.

3. **Security and Trust Validation:**
   - Tool: [*Hyperledger*]
   - Applicable KPIs: Level of Trustworthiness, Authentication Rate, Transaction Throughput, Smart Contract Execution.

- Description:Security-related KPIs will be validated using tools that track blockchain-based authentication, trust scores, and contract execution rates. These tools are integrated into SAFE-6G's security function to ensure data integrity and trustworthiness.

What is measured

Hyperledger is used to observe and validate the blockchain-based security functions implemented in SAFE-6G, with emphasis on authentication operations, transaction rates, and the correct execution of smart contracts that implement trust and policy logic.

How it is measured

- Authentication Rate is derived from counts of successful and failed authentication transactions per unit time, as found in ledger records or associated access-control logs.
- Transaction Throughput is computed from block creation logs and transaction metadata, providing metrics such as transactions per second, block size, and commit latency.
- Smart Contract Execution indicators are based on contract logs and events, which give execution times, resource usage, and success/failure ratios.
- Level of Trustworthiness is obtained from blockchain-stored credentials, reputation information, and policy-compliance events, aggregated into composite trust scores.

Integration in SAFE-6G and EOSC

Hyperledger nodes could be part of the SAFE-6G security framework. Audit trails, transaction statistics and trust indicators are exported as anonymised, structured data and deposited into EOSC-compatible services. This allows independent verification, reproducible security analysis, and cross-domain trust studies.

Note: additional tools could be identified during the progress of the project.

4. **Privacy Metrics and Governance:**
   - Tool: [*Thanos/Elasticsearch/Aeros DataFabric privacy management*]
   - Applicable KPIs: Privacy Score Calculation, Privacy Action Recommendation Latency.
   - Integration: Privacy-related KPIs are validated by tools that monitor data flow and govern access control in alignment with privacy policies. Integration with SAFE-6G's privacy module enables real-time monitoring and adjustment based on user-defined privacy settings.

What is measured

These tools track data access, consent enforcement and policy decisions along SAFE-6G data flows. They support the computation of privacy scores and the timing of automated privacy recommendations and actions.

How it is measured

- For Privacy Score Calculation, logs of data access, policy decisions, consent status and anonymisation/pseudonymisation operations are indexed in Elasticsearch and managed via Aeros DataFabric. A predefined scoring model is applied to these events, considering factors such as sensitivity of attributes, access frequency and policy compliance.
- Privacy Action Recommendation Latency is measured as the time between a privacy-relevant trigger (for instance, a new processing request or a change in user preference) and the corresponding recommendation or enforcement decision produced by the privacy engine. Timestamps are collected and aggregated through Thanos to allow temporal analysis.

Integration in SAFE-6G and EOSC

The privacy module could use Thanos for scalable, long-term time-series storage and Elasticsearch/Aeros for searchable logs and metadata. Privacy-related KPI datasets (scores, latencies, policy contexts) are exported in de-identified form and deposited into EOSC services with appropriate access controls and consent-aware metadata. This demonstrates EOSC-aligned, privacy-preserving data management.

Note: additional tools could be identified during the progress of the project.

5. **Resilience and Reliability:**
   - Tool: [*Cloud-Native agents, e.g., Cilium, Gatus*]
   - Applicable KPIs: Service Uptime, Reliability Profiling, Network Resilience.
   - Description: Resilience tools simulate network disruptions, measure uptime, and assess SAFE-6G's ability to maintain service continuity. Data from these tools feed into the resilience module, allowing proactive adjustments and performance optimization.

What is measured

These agents observe service health, connectivity and failure behaviour in SAFE-6G microservices and network paths, enabling quantitative assessment of uptime and resilience in the presence of faults.

How it is measured

- Gatus runs periodic health checks (HTTP, TCP, ICMP, or custom probes) against SAFE-6G services. Service Uptime is calculated as the proportion of successful checks over the total number in the selected evaluation period.
- Cilium provides visibility into L3–L7 connectivity and policy enforcement. By analysing connection failures, policy drops and retry behaviour, the system derives Reliability Profiling and Network Resilience metrics, including recovery times, failover success rates and error budgets.
- Fault-injection scenarios can be combined with these agents to study behaviour under controlled failures.

Integration in SAFE-6G and EOSC

Resilience-related metrics could be collected by the SAFE-6G resilience module and stored in the data fabric as time-series and event logs. Curated datasets describing failures, detection times, recovery behaviour and associated topology are then published to EOSC-compatible services, supporting comparative resilience and robustness studies.

Note: additional tools could be identified during the progress of the project.

6. **AI/ML Framework and Cognitive Coordinator KPIs:**
   - Tool: [*Kubeflow Metrics, Logging, TensorFlow, PyTorch*]
   - Applicable KPIs: AI Model Convergence Time, Cognitive Coordinator Classifier Precision, Energy Efficiency of ML Models.
   - Description: AI/ML tools provide training, testing, and deployment capabilities for machine learning models within SAFE-6G. They are integrated within the cognitive coordinator, where they assess cognitive and autonomic functions, enabling continuous model updates and accuracy validation.

What is measured

These tools support the training, validation, deployment and monitoring of AI/ML models in the SAFE-6G cognitive coordinator, with a focus on both performance and efficiency.

How it is measured

- AI Model Convergence Time is taken from Kubeflow and framework logs (TensorFlow/PyTorch) by measuring elapsed time and number of epochs or iterations until convergence criteria are met (for example, loss below a threshold or stable validation accuracy).
- Classifier Precision and related metrics (recall, F1-score, ROC-AUC) are computed from evaluation datasets and inference logs, which are tracked automatically in Kubeflow pipelines and experiment tracking components.
- Energy Efficiency of ML Models is estimated by aligning training and inference timelines (from Kubeflow) with node-level power measurements (e.g. from Kepler or similar exporters) to derive indicators such as energy per inference or per training epoch.

Integration in SAFE-6G and EOSC

The cognitive coordinator uses these metrics to manage model lifecycle and adaptation. Experiment configurations, datasets and performance results are stored in the SAFE-6G data fabric and exported to EOSC-compatible repositories as structured experiment artefacts, enabling reproducible ML studies and cross-project comparison.

Note: additional tools could be identified during the progress of the project.

7. **Energy Efficiency and Sustainability:**
   - Tool: [*Kepler*]

- Applicable KPIs: NFV Green Deployment, Energy Consumption, Environmental Sustainability.
- Description: Energy and environmental impact tools track power usage and emission metrics across SAFE-6G deployments. These tools interface with the continuums' data fabric, helping to characterize if the sustainability goals are met.

What is measured

Kepler monitors energy usage at node, container and workload level across the SAFE-6G infrastructure, with emphasis on virtualised network functions (VNFs/CNFs) and service chains.

How it is measured

- Energy Consumption is obtained from hardware counters and power models exposed by Kepler and aggregated per node, pod, or service over time.
- NFV Green Deployment metrics are derived by relating energy consumption to different deployment choices (such as placement strategies and scaling policies) and workload levels, allowing a comparison between alternative configurations.
- Environmental Sustainability indicators are calculated by combining measured energy usage with appropriate emission factors (e.g. gCO_2eq/kWh), resulting in per-service or per-scenario sustainability KPIs.

Integration in SAFE-6G and EOSC

Kepler exporters could be part of the SAFE-6G observability stack and continuums' data fabric. Energy and emission datasets, together with deployment and workload descriptors, are packaged and deposited as FAIR datasets into EOSC services, so that the environmental impact of SAFE-6G deployments can be evaluated transparently.

Note: additional tools could be identified during the progress of the project.

7.2 Integration into the SAFE-6G Platform

The SAFE-6G platform will integrates all these tools in a modular architecture. Each tool is connected to the relevant framework components, such as the cognitive coordinator, privacy module, security function, AI/ML framework, resilience module and orchestration layer. Measurement and validation tools are placed at the appropriate points in the data and control flows and within the corresponding system components to maximize their usefulness and efficiency.

By following TMV WG tool set definitions, SAFE-6G maintains compatibility with established industry standards and supports a coherent integration process. This results in a robust, validated and trustworthy 6G ecosystem in which tool placement directly contributes to performance monitoring, system adaptability and operational robustness.

To ensure that its measurements can be reused and independently analyzed, SAFE-6G will explicitly connects its KPI/KVI tool outputs to EOSC services.

- All tools will export their outputs into the SAFE-6G data fabric using open, documented formats (for example, time-series data, tabular data and structured metadata).

- Curated datasets will be deposited in EOSC-compatible repositories, accompanied by rich metadata describing the KPIs/KVIs, scenarios, topology, configuration and anonymization state, in line with FAIR principles.
- These datasets will be registered so that they can be discovered via EOSC discovery services, enabling cross-project comparison, re-analysis and validation by the broader research community.

8 CONCLUSION

The SAFE-6G project provides an advanced framework for the future of user-centric 6G networks. It highlights the importance of KPIs and KVIs in defining, monitoring, and achieving success across technical, societal, and environmental dimensions. By addressing critical challenges, SAFE-6G establishes itself as a robust and adaptive framework to meet the evolving needs of users and society. At the heart of SAFE-6G lies the LoTw metric and a dynamic, AI-driven architecture designed to adapt to user-specific requirements in real time. This architecture enables automated decision-making processes that optimize performance and ensure trustworthiness, security, privacy, resilience, and reliability. The cognitive coordination of network resources and services allows SAFE-6G to dynamically adjust operations, ensuring the delivery of high-quality, user-centric services. Such an approach not only enhances system efficiency but also strengthens trust between users and the network.

SAFE-6G places significant emphasis on aligning its KVs with global societal priorities, particularly the UN SDGs. Through its defined KVIs, the framework ensures measurable contributions to sustainability, digital inclusion, economic innovation, and environmental stewardship. These indicators are critical for evaluating the broader impact of 6G technologies, demonstrating how they can empower underserved communities, foster innovation, and contribute to a sustainable and equitable future.

The role of KPIs in the SAFE-6G framework is pivotal. They provide a clear and measurable way to monitor and evaluate the success of actions taken within the network, particularly automated ones. Automated actions, enabled by the framework's AI-driven architecture, are essential for maintaining system efficiency and responsiveness. By integrating KPI monitoring into continuous validation processes, SAFE-6G ensures that automated actions align with the intended goals of the framework. Metrics such as response time, success rates, and trustworthiness scores are evaluated in real time, allowing for rapid adjustments and optimizations. This proactive approach guarantees that automated actions remain effective and impactful, reducing risks and maximizing benefits.

The SAFE-6G framework further incorporates advanced validation methodologies that integrate AI/ML technologies and continuous integration/deployment pipelines. These methodologies enable ongoing evaluation of KPIs and KVIs, ensuring that the framework can dynamically respond to changing user needs and environmental conditions. By leveraging tools for performance testing, scalability analysis, and impact assessment, SAFE-6G ensures that its systems remain resilient, efficient, and aligned with societal priorities.

By bridging the gap between societal needs and technological advancements, SAFE-6G reinforces the importance of a user-centric approach in the design and deployment of 6G systems. Its user centric design, supported by LoTw and dynamic programmability, ensures that trustworthiness remains a cornerstone of 6G networks. The inclusion of KVIs and KPIs ensures that all aspects of the framework are measurable, traceable, and aligned with broader societal objectives.

In conclusion, SAFE-6G sets a new standard for 6G development by combining technological excellence with a commitment to societal impact. Its dynamic architecture, AI-driven automation, and robust monitoring systems demonstrate the critical role of KPIs in evaluating and sustaining success. By aligning with SDGs and addressing pressing global challenges, SAFE-6G ensures that 6G networks become a driving force for innovation, equity, and sustainability in the years to come.

9 REFERENCES

[1] TMV, from https://smart-networks.europa.eu/sns-ju-working-groups/
[2] GSMA, Mobile Industry Impact Report 2022, from https://www.gsma.com/solutions-and-impact/connectivity-for-good/external-affairs/wp-content/uploads/2022/09/sdg-main-report-2022-web.pdf
[3] 6G Flagship Research, from https://www.6gflagship.com/white-papers/
[4] 6G AI WP, 6G Artificial Intelligence White Paper, from https://6g-ia.eu/wp-content/uploads/2025/01/sustainability_of_6g-path_forward_v1.2.1.pdf
[5] HEXA-X, from https://hexa-x.eu/
[6] 6G SNS IA Research, from https://5g-ppp.eu/wp-content/uploads/2022/05/What-societal-values-will-6G-address-White-Paper-v1.0-final.pdf
[7] Kastner, «On the Relation of Trust and Explainability: Why to Engineer for Trustworthiness, » de IEEE 29th International Requirements Engineering Conference Workshops (REW), Notre Dame, IN, USA, 2021.
[8] Alexandropoulos, I., & Koumaras, H. (2025). Cognitive Coordinator Dataset of User Intent for Trustowrthiness on five principles (Reliability, Privacy, Security, Resilience, and Safety) [Data set]. Zenodo. https://doi.org/10.5281/zenodo.15512671
[9] Drampalou, S., Uzunidis, D., Vetsos, A., Karkazis, P., & Koumaras, H. (2025). An Open-Source Framework and Dataset for Multi-Layer Monitoring and Predictive Autoscaling of 6G Video Streaming [Data set]. Zenodo. https://doi.org/10.5281/zenodo.17523181

ANNEX 1: GLOSSARY OF TERMS

General Terms

- **Level of trustworthiness (LoTw):** provided to a user is the result of combined actions taken by the cognitive AI coordinator of the 6G system simultaneously at the application plane, core network plane and resource/cloud continuum plane, using the openness and programmability capabilities of these planes. By deploying Trust Functions (i.e. specialized AFs) that interface with these planes, the cognitive coordinator achieves to apply security, privacy, safety, resilience and reliability actions/measures that aim to accommodate the user's intent and realize a user-centric trustworthy service provision over 6G.
- **Trust and Trustworthiness:** Trust is an attitude that a tenant has towards a 6G system. In contrast, trustworthiness is a system property that creates trust to the 6G tenant/user. A user/tenant trusts (or requires a specific level of trust from) a 6G system, because the 6G system is trustworthy. In other words, the trustworthiness of a 6G system contributes to building the trust level of the tenant/user of the specific system. Thus, the more trustworthy the 6G system is, the higher the trust level of the tenant/user will be [7] .

6G Core Terminology

- **Service Based Architecture (SBA):** The SBA consists of new architecture where the previous 4G network functions such as Mobility Management Entity (MME), Serving Gateway (SGW), Packet Gateway (PGW) and Home Subscriber Server (HSS), Policy and Charging Rules Function (PCRF) are decomposed into more specific network functions with additional web interfaces for interconnection between them.
- **Mobile Packet Core:** Consists of all the network functions required for providing mobile network connectivity and includes but not limited to; Access Mobility Function (AMF), Session Management Function (SMF), User Plane Function (UPF), Unified Data Management (UDM), Authentication Service Function (AUSF), Unified Data Repository (UDR), Network Repository Function (NRF), Network Slice Selection Function (NSSF), Network Slice Access Control Function (NSCAF), Network Data Analytics Function (NWDAF), Network Exposure Function (NEF), Policy Control Function (PCF) Security Edge Protection Proxy (SEPP).

Openness and Programmability terminology

- **Openness:** refers to the creation of a network environment that supports interoperability, transparency, and flexibility. Open architectures and standards are designed to reduce reliance on proprietary hardware and software by promoting open interfaces, modular components, and multi-vendor ecosystems.
- **Programmability**: refers to the ability to dynamically configure, control, and customize network functions and services through software interfaces. Programmability enables 6G networks to support diverse use cases, such as IoT, autonomous vehicles, and real-time applications, by adjusting network behavior in real-time to meet unique demands.

- **Monitoring:** refers to the continuous observation and analysis of 6G network performance and user behavior, ensuring the targeted quality offered to the application users. Continuous monitoring systems can monitor the different planes and devices of the 6G ecosystem and provide metrics through frameworks, like the Common API Framework (CAPIF), to ML methods for providing estimations to achieve a specific LoTw.
- **Application Programming Interface (API):** is a set of rules or protocols that enable software applications to communicate with each other to exchange data, features and functionality.

MLOps / DataOps terminology

- **Artificial Intelligence (AI)**: Discipline to create machines with the ability to perform some cognitive functions usually associated with human minds, such as perceiving, reasoning, learning, interacting with the environment, problem solving, and even exercising creativity.
- **Machine Learning (ML):** Subfield of AI that, using data and algorithms, provides machines with the capability of learning the way humans do, without specific programming, gradually improving its accuracy.
- **Deep Learning (DL):** Type of ML that uses artificial neural networks to learn from data, as opposed to classical ML based on Decision Trees, Support Vector Machine (SVM), etc.
- **Federated/Collaborative Learning:** Type of ML that uses collaborative data provided by multiple entities but ensuring that their data remains decentralized and there is no exchange of data from client devices to global servers as the learning process takes place locally, increasing data privacy.
- **ML models:** Computer programs or systems created from ML algorithms used to recognize patterns in data or make predictions.
- **Trained ML models:** ML models that once trained can make useful predictions from new input data.
- **Supervised ML models:** ML models that use labeled data as input.
- **Unsupervised ML models:** ML models that use unlabeled data as input.
- **Reinforcement ML models:** ML models that learn by receiving feedback about performance after deployment.
- **Predictive or discriminative ML models:** ML models used for classifying or predicting data labels.
- **Generative ML models:** ML models able to generate new data / content from input data.
- **ML Operations (MLOps):** Process of managing the ML life cycle, from development to deployment.
- **Dataset**: In the context of Artificial Intelligence, a dataset is used to train learning techniques to reproduce a given behavior. We distinguish input data (provided by the system) from expected data (that we want to predict). To be used for learning, a dataset must pair inputs with expected outputs.
- **Synthetic data: Data** artificially created by computer algorithms, as opposed to real data that are collected from real events.

Chatbot terminology

- **Natural language processing (NLP):** NLP refers to the computational capability to comprehend and interpret human language. It involves recognizing patterns in communication and converting written text into spoken language.
- **Natural language understanding (NLU):** NLU is the ability of a computer system to comprehend the meanings and sentiments in natural language, facilitating a deeper understanding of human expressions.
- **Natural Language Generation (NLG)**: NLG allows the chatbot to generate human-like text responses, improving the fluidity of communication and making the interaction feel more natural and engaging.
- **Conversational AI (CAI)**: A form of artificial intelligence that powers the ability of chatbots to engage in natural language conversations with users. This allows chatbots to simulate human-like interactions, improving user experience by understanding and responding to queries in real time.
- **Sentiment analysis:** Sentiment analysis enables the evaluation of the tone and sentiment in written or spoken language. Through intent detection, chatbots and similar systems can analyze messages to assess the user's sentiment—whether positive, negative, or neutral—toward a product or service.
- **Intent detection**: Intent detection refers to the process of identifying a user's purpose or intention based on their communication.
- **Intent Classification**: Part of intent recognition, this is where the chatbot categorizes the user's input into specific predefined intents to trigger the appropriate action or response.
- **Response Generation**: The process where the chatbot formulates a response based on user input. This can involve retrieving information from databases, using AI to craft a response, or running processes like intent recognition to give meaningful answers.
- **Fallback Mechanism**: A strategy used when the chatbot fails to understand the user's intent. It provides predefined generic responses or escalates the conversation to a human agent to ensure the conversation continues smoothly.

Cognitive Coordinator terminology

- **Classification Component:** A component that uses AI algorithms, like regression, to convert textual data into numerical values for analysis and processing.
- **Queries:** Requests for specific information from a system or database.
- **Knowledge Base:** A structured database, such as MongoDB or SQL, that stores organized information for easy retrieval and analysis.
- **Reasoning Engine:** A component that retrieves relevant information from a knowledge base by executing target queries.

XAI terminology

- **Tabular data:** In this project, we will focus on so-called tabular data for inputs & outputs of AI techniques. Tabular data represents numerical or categorical values that are organized in a table. We oppose tabular data to signal data (such as images, audio files, video files…) which are unstructured. In tabular data, each input/output is structured as column for which we have a clear interpretation of the semantics.
- **Model (AI context):** In the context of Artificial Intelligence, a model is a mathematical object that can be parameterized to fit a given relationship between input data and expected outputs. Finding the right parameters for the model is computing intensive and rely on a dataset of examples that we aim to generalize. The term model encompasses both the method (various mathematical models exist, including but not limited to neural networks) and the hyperparameters learned to best fit the example dataset.
- **Explainable AI (XAI)**: XAI is a domain of Artificial Intelligence dedicated to providing explanations to end users and / or ML engineers as to the rationales behind a decision taken by a ML algorithm. This encompasses both statistical and neural net approaches.
- **Large Language Models (LLMs):** LLMs are a category of Neural Nets relying on the Transformers architecture applied to natural language understanding tasks. It is now a popular set of models to tackle human machine interaction & cooperation, including in decision taking contexts.
- **Local Explainability:** In Explainable AI, a local explanation focuses on understanding the decision of a model for a specific instance. It explains why the model made a particular prediction for an individual input.
- **Global explainability:** In Explainable AI, global explanations provide an overview of how the AI model works. It describes the overall patterns and factors that influence the model's decisions across all instances.

Security Trust Function Terminology

- **Self-Sovereign Identity (SSI):** A digital identity model where individuals have complete ownership and control over their personal data. In an SSI framework, users can store their identity information in secure digital wallets and share it selectively with service providers. This approach minimizes reliance on centralized authorities, enhancing privacy and security by preventing unauthorized access and reducing the risk of data breaches.
- **Decentralized Identifier (DID):** A globally unique identifier that enables verifiable, decentralized digital identities. DIDs are not tied to any centralized registry or authority and are often recorded on distributed ledger technologies like blockchain. They allow individuals and entities to establish secure, private connections and control their identifiers independently, facilitating self-sovereign identity management.
- **Verifiable Credentials (VC):** A tamper-evident, digital credential that can be cryptographically verified. Verifiable credentials are issued by trusted authorities and contain information about an individual or entity. They enable secure sharing of credentials (e.g., passports, diplomas,

certifications, etc.) in a way that the recipient can trust the authenticity and integrity of the information without exposing unnecessary personal data.

- **Intrusion Detection System (IDS):** A security solution designed to monitor network or system activities for malicious actions or policy violations. An IDS analyzes traffic patterns to detect suspicious activities, such as unauthorized access attempts or malware. Upon detection, it alerts administrators but does not take direct action to block the threats, serving as an early warning system for potential security incidents.
- **Intrusion Prevention System (IPS):** An advanced security mechanism that not only detects potential threats like an IDS but also takes proactive measures to prevent them. An IPS sits in line with network traffic and can automatically block or reject malicious activities in real-time by dropping packets, terminating connections, or configuring firewalls, thereby actively protecting the network from attacks.
- **Smart Contracts:** Self-executing contracts with the terms of the agreement directly written into code. Smart contracts automatically enforce and execute the agreed-upon rules and obligations when predefined conditions are met, without the need for intermediaries. They operate on blockchain platforms, ensuring transparency, immutability and security in transactions ranging from financial exchanges to supply chain management.
- **Blockchain:** A decentralized, distributed ledger technology that records transactions across a network of computers. Each block contains a list of transactions and is linked to the previous one using cryptography, forming a chain. Blockchain ensures data integrity, transparency, and security by making it virtually impossible to alter past records without consensus from the network. It is the foundational technology behind cryptocurrencies like Bitcoin and enables various applications like smart contracts and decentralized apps.

Safety Trust Function Terminology

- **Software Defined Perimeter Gateway (SDPGW):** SDPGW is the dynamically created function Gateway to redirect user traffic to specific/target applications.
- **Software Defined Perimeter Control Function (SDPCF):** SDPCF is the dynamically created function that is going to manage gateway and user connectivity lifecycle.
- **AiA:** Nested AI Agent is responsible for decision making and communicating with Cognitive Coordinator. The specific policies that are going to be applied to the UE's connectivity derive from the output of the AI agent.
- **Authentication and Authorization (AnA):** AnA is a subfunction of Safety AF, which is responsible for the proper authentication of UE/device and communicating with Network Functions such as NWDAF, UDM and PCF to collect the user information and create applicable model for requested service.
- **Function Lifecycle Management (FLCM):** FLCM is a subfunction of Safety AF, which is responsible to communicate with MANO/aerOS to request the action to deploy or remove the dynamically created SDPGW & SDPCF.

Privacy Trust Function Terminology

- **Privacy:** Privacy refers to the protection of personal data and communication information from unauthorized access or misuse in 5G networks and services.
- **Level of Privacy (LoP):** The **LoP** that can be guaranteed to the specific service at a given time.
- **Desired Level of Privacy (dLoP):** The **LoP** that the **Cognitive Coordinator** requested for a specific service.
- **Final Level of Privacy (fLoP):** The maximum **LoP** that can be achieved by the network at a given time for a given service. It is only calculated and reported to the **Cognitive Coordinator** when the **dLoP** cannot be satisfied.
- **Privacy Action:** An action that can be evoked by the **Privacy Function** and affects the **LoP** of the given service.
- **Decision Support System (DSS):** The **DSS** is responsible for deciding the **Privacy Actions** that need to be taken to achieve a higher **LoP** for the service.

Resilience Trust Function Terminology

- **Resilience**: refers to the capability of the network to sustain service quality and reliability through efficient resource management, customizing them from user-intents and adapting to variations in network loads and infrastructure conditions to ensure continuous service availability.
- **Level of Resilience (LoR)**: refers to the resilience level that can be provided to a specific service at a given moment, based on current network resource allocation and infrastructure capacity.
- **Desired Level of Resilience (dLoR)**: refers to the resilience level requested by the Cognitive Coordinator based on an input from LoTw for a service to handle expected changes in network resource demand and infrastructure load.
- **Resilience Action:** refers to an action triggered by the Resilience Function to apply, adjust or improve the LoR of a user for a specific service, such as prioritizing, rerouting traffic or allocating additional resources.

Reliability Trust Function Terminology

- **Reliability:** refers to the ability of the 6G network to consistently perform its intended function under predefined conditions, ensuring the targeted quality offered to the application users. Various reliability-related mechanisms are offered by the SAFE-6G to provide a high LoTw, including data collection from different planes and devices, training, and deployment of ML methods, as well as generation of alerts.
- **Alerts:** refer to mechanisms designed to disseminate critical information quickly and efficiently during cases of significant events. Examples of significant events include Quality of Experience (QoE) degradation, reaching system breaking points, detection of abnormal operations, etc. The alerts leverage the capabilities of SAFE-6G to ensure the LoTw.
- **Service Profiling** refers to the process of defining, analyzing, and optimizing the diverse services supported by the 6G network. This involves analyzing and understanding the specific

requirements of different applications and user scenarios to ensure that the network can deliver the necessary performance, reliability, quality of service (QoS) and QoE to the application users. The service profiling aids in identifying patterns and potential bottlenecks that could negatively impact reliability. Service profiling also aids in assessing how innovative technologies might introduce vulnerabilities to the 6G system, allowing for taking preemptive measures.