**Grant Agreement No.:** 101096342
**Call:** HORIZON-JU-SNS-2022
**Topic:** HORIZON-JU-SNS-2022-STREAM-B-01-04
**Type of action:** HORIZON-JU-RIA

Holistic, omnipresent, resilient services
for future 6G wireless and computing ecosystems

# D2.1 HORSE Landscape: Technologies, State of the Art, AI Policies and Requirements

Revision: v.1.0

| | |
|---|---|
| Work package | WP2 |
| Task | T2.1 |
| Due date | 30/06/2023 |
| Submission date | 30/06/2023 |
| Deliverable lead | 8BELLS |
| Version | 1.0 |
| Authors | Ioannis Siachos (8Bells), Georgios Kontopoulos (8Bells), Vito Cianchini (MARTEL), Massimo Neri (MARTEL), Paulo Paixão (EFACEC), Pedro Elísio (EFACEC), Orazio Toscano (ETI), Stefanos Venios (S5), Jose Manuel Manjón (TID), Carina Pamminger (HOLO), Clayton Gordy (HOLO), Iulisloi Zacarias (TUBS), Panagiotis Gkonis (NKUA), Eva Rodriguez (UPC), Jose Manuel Manjón (TID), Sofia Giannakidou (STS), Fabrizio Granelli (CNIT) |
| Reviewers | Vito Cianchini (MARTEL), Massimo Neri (MARTEL) |

| | |
|---|---|
| Abstract | *D2.1 HORSE Landscape: Technologies, state of the art, AI policies and requirements* is a public report that sets the ground for the technical work to be done in the HORSE project. First, a State of the Art analysis in 6G, Artificial Intelligence and cybersecurity is presented. Then the two HORSE use cases are analysed in the context of the project's architecture. Additionally, the platform's functional and non-functional requirements are specified. Also, some 6G services and threats that feed the design of HORSE are discussed. Lastly, the data management procedures of the Artificial Intelligence algorithms to be used are also formalized. |
| Keywords | State of the Art, Requirements, Policies, 6G, Artificial Intelligence, Cybersecurity, Extender Reality (XR), Light Rail Transit (LRT) |

## DOCUMENT REVISION HISTORY

| Version | Date | Description of change | List of contributor(s) |
|---|---|---|---|
| V0.1 | 01/03/2023 | 1st version of the template for comments | Miguel Alarcón (Martel) |
| V0.2 | 10/03/2023 | SotA topics fixed and ToC added | Ioannis Siachos (8Bells) |
| V0.3 | 31/03/2023 | Initial UC descriptions and preliminary requirements. Merging with SotA. | Iulisloi Zacarias (TUBS) / Eva Rodriguez (UPC)/ Josep Martrat (ATOS) / Ioannis Siachos (8Bells) |
| V0.4 | 12/04/2023 | Content after Use Case workshop for requierement elicitation | UC leaders (EFACEC / HOLO) |
| V0.5 | 28/04/2023 | Several updates SotA and R&D projects. | ALL partners and Ioannis Siachos (8Bells) consolidate |
| V0.6 | 16/05/2023 | Content as a result of B5G & Security internal workshop | NKUA, ATOS, SUITE5, ETI, TID, UPC |
| V0.7 | 31/05/2023 | Refined SotA | ATOS, NKUA, 8BELLS, CNIT |
| V0.8 | 12/06/2023 | Update list of requirements from UCs | Iulisloi Zacarias (TUBS) and other partners at General Assembly in Vilanova |
| V0.81 | 19/06/2023 | Update list of requirements based on discussions in GA meeting in Vilanova | Iulisloi Zacarias (TUBS) |
| V0.9 | 23/06/2023 | Update after GA meeting in Vilanova and version for internal review. | ALL |
| V0.91 | 26/06/2023 | Internal review. Added Executive Summary. | 8BELLS, MARTEL, ATOS |

| | | | |
|---|---|---|---|
| V1.0 | 30/06/2023 | Update with review. QA and submission | 8BELLS, ATOS, CNIT |

## Disclaimer

Co-funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the other granting authorities. Neither the European Union nor the granting authorities can be held responsible for them.

## Copyright notice

| Project co-funded by the European Commission in the Horizon Europe Programme | | |
|---|---|---|
| **Nature of the deliverable:** | R | |
| **Dissemination Level** | | |
| **PU** | *Public, fully open, e.g. web* | **x** |
| **SEN** | *Sensitive, limited under the conditions of the Grant Agreement* | |
| **Classified R-UE/ EU-R** | *EU RESTRICTED under the Commission Decision No2015/ 444* | |
| **Classified C-UE/ EU-C** | *EU CONFIDENTIAL under the Commission Decision No2015/ 444* | |
| **Classified S-UE/ EU-S** | *EU SECRET under the Commission Decision No2015/ 444* | |

* *R: Document, report (excluding the periodic and final reports)*
*DEM: Demonstrator, pilot, prototype, plan designs*
*DEC: Websites, patents filing, press & media actions, videos, etc.*
*DATA: Data sets, microdata, etc*
*DMP: Data management plan*
*ETHICS: Deliverables related to ethics issues.*
*SECURITY: Deliverables related to security issues*
*OTHER: Software, technical diagram, algorithms, models, etc.*

# Executive summary

This document provides an overview of the HORSE project’s vision, research pillars, use cases, functional and non-functional requirements, network services and threats, AI data management, and conclusions.

The document begins with an introduction that outlines the purpose, methodology, document structure, and relation to other work packages and tasks. It then delves into the HORSE vision and background, highlighting its mission and underlying technologies.

The research pillars of HORSE are discussed next, with a focus on security, networking, and artificial intelligence. In the security domain, the document explores the implications of 6G networks, including new technologies/architectures, physical layer security, privacy protection, and risks and threats. Threat identification, characterization, and modelling techniques such as anomaly detection and threat modelling are also examined.

In the networking domain, the document addresses network exposure capabilities beyond 5G, energy efficiency, digital twin design, and the role of the physical layer in 6G networks. Furthermore, it explores AI-enabled solutions for security enhancement and threat mitigation, as well as intent-based networking.

The HORSE use cases section presents two specific scenarios: Secure Smart LRT Systems (SS-LRT) and Remote Rendering to Power XR Industrial (R22XRI). These use cases are described, their problem statements are analysed in relation to the HORSE infrastructure, and various usage scenarios are demonstrated.

The functional and non-functional requirements of HORSE are outlined, covering the necessary capabilities and characteristics of the system. Additionally, the document discusses network services considered in HORSE, including network slicing isolation, API exposure, and AI data training. It also identifies the threats associated with the use cases, specifically in secure smart LRT systems and remote rendering for XR industrial applications.

HORSE AI data management is explored, focusing on data ingestion mechanisms, data management procedures, and description of the datasets used in the use cases. The document provides details on the datasets related to the secure smart LRT systems and XR industrial use cases, as well as 5G/6G network traffic data.

In conclusion, the document summarizes the key findings and insights from the HORSE project, emphasizing the importance of addressing security, networking, and AI challenges in the context of 6G networks.

The references and an appendix listing EU beyond-5G and 6G research projects are provided for further exploration.

# Table of contents

# List of figures

Co-funded by the European Union

# List of tables

# Abbreviations

| | |
|---|---|
| **AF** | Application Function |
| **AI** | Artificial Intelligence |
| **AMF** | Access and Mobility Management Function |
| **API** | Application Programming Interface |
| **AUSF** | Authentication Server Function |
| **CAD** | Computer Aided Design |
| **CAPIF** | Common API Framework |
| **CC** | Command Centre |
| **CCTV** | Closed-Circuit TV |
| **COTS** | Commercial Of The Shelf |
| **CPS** | Cyber-Physical Systems |
| **c-RAN** | cloud-RAN |
| **CVSS** | Common Vulnerability Scoring System |
| **D2D** | Device-to-Device |
| **DL** | Deep Learning |
| **DoS** | Denial of Service |
| **DRL** | Deep Reinforcement Learning |
| **DT** | Digital Twin |
| **Dx.x** | Deliverable x.x |
| **ECC** | Elliptic Curve Cryptography |
| **FDL** | Federated Deep Learning |
| **FL** | Federated Learning |
| **FSO** | Free Space Optics |
| **GAN** | Generative Adversarial Network |
| **IBN** | Intent-Based Networking |
| **IDS** | Intrusion Detection Systems |
| **IIT** | Industrial Internet of Things |
| **IoT** | Internet of Things |
| **IT** | Information Technology |
| **KPI** | Key Performance Indicator |
| **LRT** | Light Rail Transit |

| | |
|---|---|
| **LSTM** | Long-Short Term Memory |
| **MEC** | Multi-Access Edge Computing |
| **MIMO** | Multiple-Input Multiple-Output |
| **MISO** | Multiple-Input Single-Output |
| **MITM** | Man in the Middle |
| **ML** | Machine Learning |
| **Mxx** | Month xx |
| **NEF** | Network Exposure Function |
| **NFV** | Network Function Virtualisation |
| **NOMA** | Non-Orthogonal Multiple Access |
| **NRF** | NF Repository Function |
| **NSSF** | Network Slice Selection Function |
| **NWDAF** | Network Data Analytics Function |
| **OCC** | Operational Command Centre |
| **OT** | Operational Technology |
| **PCF** | Policy Control Function |
| **PID** | Public Information Devices |
| **PKI** | Public Key Infrastructure |
| **PLS** | Physical Layer Security |
| **R2XRI** | Remote Rendering to Power XR Industrial |
| **RBM** | Restricted Boltzmann Machine |
| **RIS** | Reconfigurable Intelligent Surfaces |
| **RNN** | Recurrent Neural Networks |
| **SBA** | Service-Based Architectures |
| **SDN** | Software Defined Networking |
| **SMF** | Session Management Function |
| **SRad** | Symbiotic Radio |
| **SS-LRT** | Secure Smart LRT Systems |
| **TL** | Transfer Learning |
| **Tx.x** | Task x.x |
| **UAVs** | Unmanned Aerial Vehicle |
| **UDM** | Unified Data Management |
| **UPF** | User plane function |

| | |
|---|---|
| **V2X** | Vehicle-to-Everything |
| **VLC** | Visible Light Communications |
| **VM** | Virtual Machine |
| **VR** | Virtual Reality |
| **WPx** | Work Package x |
| **XR** | Extended Reality |

Co-funded by the European Union

# 1. Introduction

## 1.1. Purpose of this Document

This document serves as a comprehensive report on the entire HORSE context, specifically focusing on the IT-1 architectural design. Its purpose is to provide a detailed overview and analysis of the technologies, state of the art, AI policies, and requirements relevant to HORSE. By consolidating information from various sources and leveraging the expertise of multiple partners, this document aims to inform and support the decision-making process, guiding the architectural design phase of HORSE. It offers a thorough understanding of the landscape surrounding HORSE, enabling stakeholders, researchers, policymakers, and practitioners to make informed decisions and recommendations, ensuring the proper implementation of the project.

## 1.2. Methodology

The development of this document has been a collaborative effort involving multiple partners. The methodology employed in its creation aimed to ensure a comprehensive and systematic analysis of the landscape of technologies, current state of the art, AI policies, and requirements in the context of HORSE.

The methodology can be outlined as follows:

- **Requirements Gathering:** The partners collectively identified the key objectives of the document and defined the requirements for each section. This included determining the scope, purpose, and target audience of the landscape analysis, as well as the specific aspects to be covered. A thorough literature review was conducted to identify existing research, reports, policies, and industry standards related to the technologies, state of the art, AI policies, and requirements relevant to HORSE. This served as a foundation for the subsequent analysis.
- **Data Collection and Analysis:** The partners collaborated in gathering data from various sources, including academic research papers, industry publications, government regulations, and relevant online resources. The data collected encompassed a broad spectrum of topics related to the landscape of technologies, state of the art, AI policies, and requirements. A systematic analysis of the collected data was conducted, involving categorization, synthesis, and comparison. The partners utilized their expertise and collective knowledge to identify trends, patterns, gaps, and opportunities within the field.
- **Document Structure and Writing:** Based on the outcomes of the analysis, a structured outline for the document was developed. This ensured a logical flow of information and facilitated the presentation of findings and recommendations. The partners collaborated on writing the content for each section, drawing upon their respective expertise and insights. Rigorous review and feedback processes were implemented to ensure accuracy, coherence, and clarity of the document.
- **Review and Validation:** The draft document underwent several rounds of review by all partners. Feedback and suggestions were incorporated, and revisions were made to enhance the document's quality and comprehensiveness. The final version of the document represents the collective knowledge, expertise, and consensus of the partner consortium.

By following this methodology, the partners aimed to provide a robust and reliable resource that informs stakeholders, researchers, policymakers, and practitioners about the current landscape and future directions in the field of HORSE.

## 1.3. Document Structure

This document is structured into several sections to provide a comprehensive understanding of the HORSE landscape, technologies, state-of-the-art analysis, AI policies, and requirements. The following is an overview of the document's structure:

- **Section 1: Introduction:** This section introduces the document, outlining its purpose, methodology, relation to other work packages and tasks, and the structure of the document itself.
- **Section 2: HORSE Vision and Background:** In this section, the HORSE mission, underlying technologies, and the consortium's shared vision for HORSE are presented.
- **Section 3: HORSE Research Pillars and State-of-the-art analysis:** This section focuses on the three main research pillars of HORSE: Security, Networking, and AI. Each pillar includes a state-of-the-art analysis conducted by specific partners within the consortium, highlighting relevant advancements and insights.
- **Section 4: HORSE Use Cases:** Here, two specific use cases within the HORSE framework are presented: Secure Smart LRT Systems (SS-LRT) and Remote Rendering to Power XR Industrial (R2XRI). Each use case includes a description, adaptation to the HORSE infrastructure, existing data models and workflows, as well as demonstration usage scenarios.
- **Section 5: HORSE Functional & Non-Functional Requirements:** This section discusses the functional and non-functional requirements of the HORSE project, focusing on aspects such as security, networking, and AI. These requirements serve as guidelines for the architectural design phase.
- **Section 6: HORSE AI Policies:** Here, the AI policies related to the HORSE project are addressed, covering topics such as ethical considerations, data protection, and governance.
- **Section 7: Takeaways for the Architectural Design:** This section highlights key takeaways and insights from the preceding sections, providing valuable guidance for the architectural design phase of HORSE.
- **Section 8: Conclusions:** The document concludes by summarizing the main findings, emphasizing the importance of the HORSE project, and providing an outlook for future developments.

By following this structure, the document aims to deliver a comprehensive analysis and guide the architectural design process within the HORSE project.

## 1.4. Relation to other Work Packages and Tasks

This Deliverable is the 1st iteration of the output of tasks *T2.1 – Market radar and baseline technologies identification, T2.2 – Overall requirements specification and identification and T2.3 – AI data collection strategy and procedures*. They all start in M01 and end in M21. Sections 2 and 3 are associated with T2.1, Sections 4 and 5 with T2.2 and Section 6 with T2.3. The report feeds task *T2.4 – Architectural design,* that with an internal report in M04 will lead to the final functional design of the HORSE platform, to be later developed in WP3 and WP4, and finally integrated and validated in WP5. This deliverable will be updated in its 2nd iteration in M21.

# 2. HORSE Vision and Background

## 2.1. Vision

The consortium's shared vision for HORSE is to create a powerful and advanced infrastructure that drives the development of novel services in the context of 6G networks. This infrastructure, referred to as the HORSE platform, aims to be human-centric, open-source, green, sustainable, and capable of seamlessly incorporating advancements in various domains. The platform will address the grand challenge of operating 6G infrastructure for smart connectivity and service management while emphasizing effectiveness at the intersection of 6G connectivity, computing infrastructure management, and security.

Cellular mobile communications have been through five generations. With the adoption of a new generation of cellular mobile communications, there are advancements in frequency, bandwidth and data rates. Currently, we are in the **5th generation (5G)** which achieves sub-6 GHz and millimetre-wave bands while maintaining a peak rate of 20 Gbps. Even though 5G communication systems provide notable advancements compared to previous generations, they still have some restrictions. Some applications and services require superior communication performance beyond the capabilities of 5G. These include global coverage, very low latency, extremely fast data transmission rates, densely connected networks, precise positioning, reliable and secure connections, low power consumption, high energy efficiency, and widespread intelligence.

While **6G** is expected to surpass these limitations, its standardization is still a work in progress [1]. It is expected for 6G to enhance 5G with the addition of AI and big data, transforming it to an intelligent network. Thus, 6G will go beyond communications, in contrast with its predecessors. To achieve global coverage, 6G will integrate ground, space, air and sea communication networks. It will fully utilize all spectra such as sub-6 GHz, THz, mmWave, and optical bands. 6G will provide users with an enhanced sensory experience with holographic communications, extended reality (XR), and other applications. With digital twin technology, 6G will map the digital world to the physical world and realize intelligent connections between humans, objects and machines. Strong security will be ensured through endogenous security, which includes physical and network layer security, and AI will be used to achieve intelligent protection [2].

HORSE project aims to create an advanced and versatile platform that leverages the ongoing evolution of 6G capabilities. It seeks to address the challenges of technology solutions and system evaluation that are not yet foreseen, ultimately enabling an omnipresent, smart, and secure network service provisioning in the future network-of-networks landscape.

The envisioned HORSE platform will be validated through two representative scenarios: distributed operation of transport systems and multiuser remote rendering in extended reality. These scenarios, provided by experienced actors in their respective fields, will demonstrate the adaptability of the HORSE platform to different constraints and requirements.

Finally, as an additional note, we refer the interested reader to Appendix A, in which a list of relevant projects that can inspire the design of the HORSE platform are presented.

## 2.2. Underlying Technologies

HORSE is a research project with very high complexity that utilizes various technologies from a vast range of research topics that come from the academic disciplines of Cybersecurity, Artificial Intelligence and Networking. As such, this section will present the main underlying technologies and concepts that are required for the understanding of HORSE's vision.

The HORSE platform will provide a coordinated provisioning and protection evolutionary platform, combining technologies such as predictive threats detection and impact analysis, proactive mitigation actions against threats and breaches, programmable networking, semantic communications, Network Function Virtualisation (NFV), intent-based networking, AI-based techniques, in-network computing, and cross-layer management of physical layer features. These technologies will be incorporated into the platform as they emerge in the 6G realm.

In the upcoming 6G network age, **Service-Based Architectures (SBAs)** are poised to be crucial to service delivery. The purpose of SBAs is to deliver services to end users through a flexible and modular framework. This design's ability to separate the service logic from the underlying network infrastructure enables quick deployment and scalability of highly customized and personalized services. This modular strategy enables network operators to respond swiftly to shifting user demands and service requirements while constructing a more dynamic and adaptable network that can serve a wide variety of use cases and applications [3]. The future's immensely adaptable and flexible 6G networks will essentially be made possible by SBAs.

**Network Function Virtualization (NFV)** is the separation of network functions from the hardware that supports them. In order to not have hard requirements in terms of hardware, they can be virtualized and run on either common servers or the cloud. As a result, the implementation of network services and functions can be more flexible, and scalable. Furthermore, NFV gives network operators the ability to scale resources, develop network services quickly, and dynamically assign resources in response to shifting consumer demand [4]. This is why the successful delivery of 6G networks will be significantly impacted by NFV.

**Multi-Access Edge Computing (MEC)** is a computing paradigm that brings computation and data storage closer to the network edge, enabling faster and more efficient data processing. MEC in 6G networks is expected to play a crucial role in supporting various applications and services, including those requiring low latency, high bandwidth, and massive device connectivity. MEC security in 6G refers to the measures and mechanisms put in place to protect the MEC infrastructure, applications, and data from potential threats and vulnerabilities. As 6G networks are anticipated to connect a vast number of devices and support critical services, ensuring robust security in MEC is of paramount importance.

**Artificial Intelligence (AI)** can be used to identify, prevent, and respond to intrusions more effectively than traditional security measures for the sophisticated **cybersecurity** requirements of 6G networks. Particularly, AI systems are able to identify trends and anomalies in network traffic, predict potential attacks, and take preventative action before they can do any harm. AI can also be used to scan vast amounts of data and spot intricate attack patterns that are challenging for human operators to recognize. Implementing AI for 6G cybersecurity faces a number of difficulties, including the necessity for substantial training datasets and the possibility of adversarial attacks. All things considered, AI is a fascinating and promising technology for the cyber protection of 6g networks.

**Digital Twins** are virtual representations of real-world systems or physical objects that can implement ongoing monitoring, simulation, and improvement. They are anticipated to be crucial in the creation and implementation of intricate and connected 6G network infrastructures. They can simulate various network components, such as antennas, base stations, and devices and increase the effectiveness and performance as a result. Also, they can be used as sandbox environments for the testing of new network services and components. However, there are many obstacles to overcome when developing digital twins for 6G networks, including the requirement for high-fidelity models, effective simulation methods, and scalable infrastructure [5].

# 3. HORSE Research Pillars and State-of-the-art analysis

HORSE, a highly complex research endeavour, harnesses a multitude of cutting-edge technologies sourced from three main research domains, specifically Cybersecurity, Artificial Intelligence, and Networking, to achieve its ambitious objectives. This interdisciplinary approach empowers HORSE to tackle intricate challenges at the intersection of these fields, opening new avenues for groundbreaking advancements. Those main research areas are being presented in a state-of-the-art analysis that focuses on 5G/6G networks in the following sections. This analysis feeds the whole design of HORSE as it's output directly influences the requirements, policies and use cases.

## 3.1. Security

### 3.1.1. Security in the 6G world

Security and privacy were already an integral pillar of the 5G architecture, and it will keep growing in importance as the evolution of mobile networks continues, and 6G will not be an exception. It is clear that security and privacy will be of a paramount importance in 6G, and in fact, it will be a critical factor for 6G success. However, security will have different facets in 6G and consequent generations of mobile networks. 6G network will be closer to humans, blurring the line between the physical and digital worlds, enabling for instance human-centric mobile networking or the precise location of a person in a room. As such, a security incident in this context could lead to a loss of information, loss of control over your devices, loss of money, loss of property, or even physical danger to people. This potential impact on safety makes security absolutely critical.

Furthermore, it is envisaged that the size of 6G cells will decrease from small cells to "tiny cells" [6]. This will have impact on the density of cells deployments, where Device-to-Device (D2D) communications, mesh networks and multi-connectivity will become the norm. In this hyper-connected and heterogeneous context, malicious devices/actors will have an expanded attack surface, with more connected and dispersed devices providing greater potential for attackers. To properly address security in these circumstances will also require a shift in the security paradigm, moving to a more proactive and decentralized approach, seeking to maximize the automation of the detection and response through the embedded use of AI.

Besides the security issues inherited from 5G, related to the virtualization of the network and services (NFV/SDN), new security and privacy challenges need to be addressed in this context to ensure the conception of a secure and trustworthy 6G network. These challenges are diverse and multidisciplinary, not limited only to technology, but also including regulation, techno-economics, politics, and ethics. They can be categorized in the following four main groups described in the subsections bellow.

#### *3.1.1.1. Trusted networking*

The future 6G network is envisioned as an ultra-large-scale network, more open and heterogeneous than previous generations, with an extremely high demand on the scalability, connecting a variety of heterogeneous resources including devices, types of networks, cloud deployments and operators with different management architectures and signalling systems. To satisfy these requirements, trust management, across multiple operators and domains, including the adoption of zero-trust technologies, must be embedded in the architecture for achieving the expected security level in 6G. This represents a disaggregated architecture with multiple multi-vendor trust domains across the cloud stack and topology, as well as untrusted domains on a massive scale consisting of sub-networks and devices. Consequently, trust modelling, trust policies and trust mechanisms need to be defined and standardized at least for the tangible aspects of trust, including common and agreed practices for establishing trust in different service scenarios and, how to measure it.

#### *3.1.1.2. New technologies/architectures*

This group of challenges includes the security and privacy concerns associated with the introduction of new technologies, which can be seen as enablers of the 6G networks, but also can enable attackers to carry out more powerful and complex attacks. Three technologies will have the greatest impact on the threat landscape in this respect: Quantum computing, AI and Cloud computing offloading.

The 5G security transport layer relies on traditional cryptography such as ECC and it is expected that, although the security architecture of 6G will be more complex, it will be still based on current transport layer security standards in which PKI will play a central role. However, the cryptographic primitives that form the basis of these security standards can be easily broken in polynomial time on a quantum computer. According to McKinsey estimation [7], quantum technology will be operational available in ten to twenty years' time, and will represent a serious threat to security for all organizations and consequently, also for 6G security.

Today, there are already some cryptographic primitives (e.g., McEliece [8] and lattice-based NTRU [9]) that are considered quantum-safe, however, their efficiency is poor, and their key sizes are large compared to traditional current schemes.

The adoption of these quantum-safe cryptographic primitives in the 6G architecture will have a negative impact on both, communications and efficiency of the network. Consequently, additional research is needed before these primitives are ready to the intended performance and functionality of the 6G architecture. In this regard, some initial work has been already done by NIST [10] identifying the potential candidates, splitting them into groups and describing the challenges they are trying to address. 3GPP [11] also analyzed the potential application of these primitives into the 5G network.

6G is expected to be something more than just a faster version of 5G, more specifically is expected to be AI-empowered. The use of AI in 6G will not be limited to its architecture (as done in 5G) but also with a deeper integration of currently emerging AI tools and networking functions, including security. Additionally, to securely support the expected bandwidth, with greater densification and cloudification in a hyper-connected world with billions of devices and nodes, will require a drastic change on the security parading, making it more proactive and automated and for this, the concepts of softwarization and virtualization in combination with AI will be crucial to make 6G secure. To overcome the security constraints existing in 5G networks and match the needs of 6G security, SDN and NFV concepts need to be further enhanced with embedded intelligence provided by AI to ensure the required dynamicity, agility and end-to-end security. Furthermore, the combination of AI/ML with edge computing will be the enabler of an intelligent edge security in 6G, fully automated, zero-touch and zero-trust security, where the network can be seen as a giant firewall that integrates the flowing security functions.6G is also expected to support a significantly increasing number of devices, and given the growing computing demands and performance limitations of typical devices, the offloading of computing, storage and networking functions to other nodes is expected to continue in 6G. This remote offloading of functions is an outstanding feature of 6G, but unfortunately brings new security concerns related to the protection of the processed data, such as:

- **Data confidentiality:** how can the data owner be sure that the data will only be used for the intended purpose?

Co-funded by the European Union

- **Data integrity:** how can the data owner be sure that the data have not been tampered during the process?
- **Platform integrity:** how to ensure that the platform itself is trustworthy in an ever-changing environment?
- **User privacy:** In addition to data confidentiality and privacy, how to ensure that user behavior has not been tracked?

#### *3.1.1.3. Physical layer security*

As mentioned previously, the expected proximity to the human' body and the implications that a security breach may have on safety makes the security of the physical layer of unprecedented relevance in 6G. The security of the physical layer has to be considered as an enabler for ensuring security and privacy in 6G, which at the same time imposes several challenges. How to address security in resource and latency constrained environments is one of them. To address this, it is envisaged that more distributed and cooperative security mechanisms will be required for threat detection, based on AI/ML models trained with massively aggregated attributes, which will be later deployed on the different security controls distributed across the network. Alternatively, the knowledge acquired in some nodes can be shared with the others by means of retrained models using federated learning. Other challenges to be addressed in this area are the prevention of attacks against cell-free MIMO and intelligent reflective surfaces, and the protection against jamming attacks.

#### *3.1.1.4. Privacy protection*

The expected hyper-connectivity to be provided by 6G will foster the creation of new and innovative smart services that will make massive use of the data collected by the myriad of devices and sensors. This will create an unprecedented opportunity for citizens, cities, governments, and industries which in turn, will rely on the sharing of large amounts of personal and/or confidential data. How this data is used, by whom and for what purposes raises several concerns about privacy leading operators and service providers to consider new governance paradigms to support the requirements imposed by the corresponding regulations.

In this context, new challenges will emerge concerning for instance the definition and quantification of deidentified data, the development of mechanisms for measuring the level of personal information present on a concrete dataset, or how to measure the required levels of protection to be required or currently provided. In this context, cyber-resilience at the core of the network will be essential to guarantee privacy and enable trust. However, for accomplishing a privacy-preserving and trustworthy 6G network in the expected multi- stakeholder and service provider environment in which the network infrastructure will be shared, new approaches for privacy protection will be required to strike a balance between preserving consumer privacy and trust.

### 3.1.2. Risks and threats

6G networks promise unprecedented advancements in connectivity, speed, and data processing. However, as the technology's capabilities increase, the system's attack surface increases and so do the potential security risks and threats.

At a high level, several macro-risks can be identified. Firstly, data privacy becomes a concern as the amount of personal data transferred and stored will increase exponentially with the anticipated rise in Internet of Things (IoT) devices, smart city infrastructures, and AI integrations. If not appropriately secured, this data could be vulnerable to theft and misuse. Secondly, as network complexity increases, it presents more points of vulnerability for attackers to exploit. The expected ubiquity of 6G networks, where virtually everything is connected, significantly raises the stakes. A single breach could disrupt entire systems or infrastructures. Moreover, with 6G's reliance on artificial intelligence for network optimization and management, malicious actors could use AI to conduct sophisticated, targeted attacks. This introduces a new level of risk and threat. Additionally, the increased use of small cells for coverage in 6G networks could lead to physical tampering. Supply chain vulnerabilities could also provide an avenue for inserting malicious hardware or software into the 6G infrastructure. These aspects raise concerns regarding the security of the network[12]–[14].

Further risks and threats can be identified and organized based on different criteria. Technology-related risks and threats encompass specific risks associated with various technologies integrated into 6G networks.

For example, AI-related risks include poisonous attacks, evasion attacks, model extraction, and model inversion attacks. Visible Light Communication (VLC) is vulnerable to eavesdropping and jamming. Terahertz technology faces access control attacks and eavesdropping. Blockchain is at risk of Sybil attacks, re-entrance attacks, and privacy attacks. Quantum communication can be targeted through quantum cloning attacks and quantum collision attacks. Molecular communication faces flooding (DoS) attacks, jamming, desynchronization, and collision attacks.

The risks and threats can also be categorized based on specific architecture layers. Sensing layer risks include physical attacks, theft of information, attacks on visible light communications, and sniffing attacks. Edge layer attacks involve data poisoning, evasion attacks, and privacy infractions. Control layer attacks target SDN, cloud computing services, and ML models. Application layer attacks pose risks in intelligent network management, such as DoS and Man-in-the-Middle (MITM) attacks, and unauthorized access to systems through intent-based interfaces.

Furthermore, application-related risks and threats arise from specific applications in 6G networks. UAVs are susceptible to physical attacks, spoofing, eavesdropping, DoS, and hijacking attacks. Holographic applications face unsecured data transmission and privacy concerns. Extended reality is vulnerable to security issues related to sharing personal data, data leakage, and unauthorized access to confidential information. Connected autonomous vehicles are at risk of capturing sensor data, physical hijacking, falsifying cloud service data, and confidentiality threats. Digital twins can be tampered with or intercepted, compromising privacy, and IoT information can be altered, infringing upon system privacy. Cyber-physical systems (CPS) face unauthorized access, data breaches, manipulation of control systems, and privacy violations.

Lastly, open RAN security risks involve insufficient isolation, privacy breaches, misconfiguration, supply chain risks, and increased opportunities for attackers. These risks and threats highlight the challenges that need to be addressed to ensure the security and resilience of 6G networks. Effective security measures, protocols, and standards must be developed and implemented to mitigate these risks and protect the integrity, privacy, and functionality of the network and its applications [15], [16].

### 3.1.3. Threats identification, characterization, and modelling

The cybersecurity threat landscape is becoming more and more complex. Threat actors are using coordinated attacks taking advantage of network security flaws to launch sophisticated attacks that have the potential to paralyze entire networks. It is therefore of paramount importance the ability to detect them rapidly and accurately and to this aim massive data streams available from diverse sources are constantly analysed for anomalous behaviours.

Two approaches to network anomaly detection are widespread used. The traditional one is the so called “signature-based” method, that detects network anomalies by looking for patterns that match signature of known anomalies. A rich literature is available about these methods and their many implementations currently used [17], [18]. The main weakness of this method is the necessity that the anomaly signatures be known in advance making it clearly not applicable for new anomalies. Moreover, the malicious attackers often workaround signature-based detection systems by garbling the signatures. The second approach is the “statistic-based” one, which does not need any prior knowledge of the nature and properties of anomalies and is, therefore, very effective even for new anomalies or variants of existing anomalies. In this context we are seeing an increasing and effective application of AI/ML solutions [19] able to establish a notion of “normality” from the training datasets and, thereafter, detect anomalies (outliers) as deviations from this normality. Taking as a reference the Cyber Threat Taxonomy Adapted from the European CSIRT [20], please refer to the following Figure 1, ML are enlarging the application scope in which they are successfully applied spanning from intrusion detection, to malware analysis, from vulnerability prioritization to spam detection.

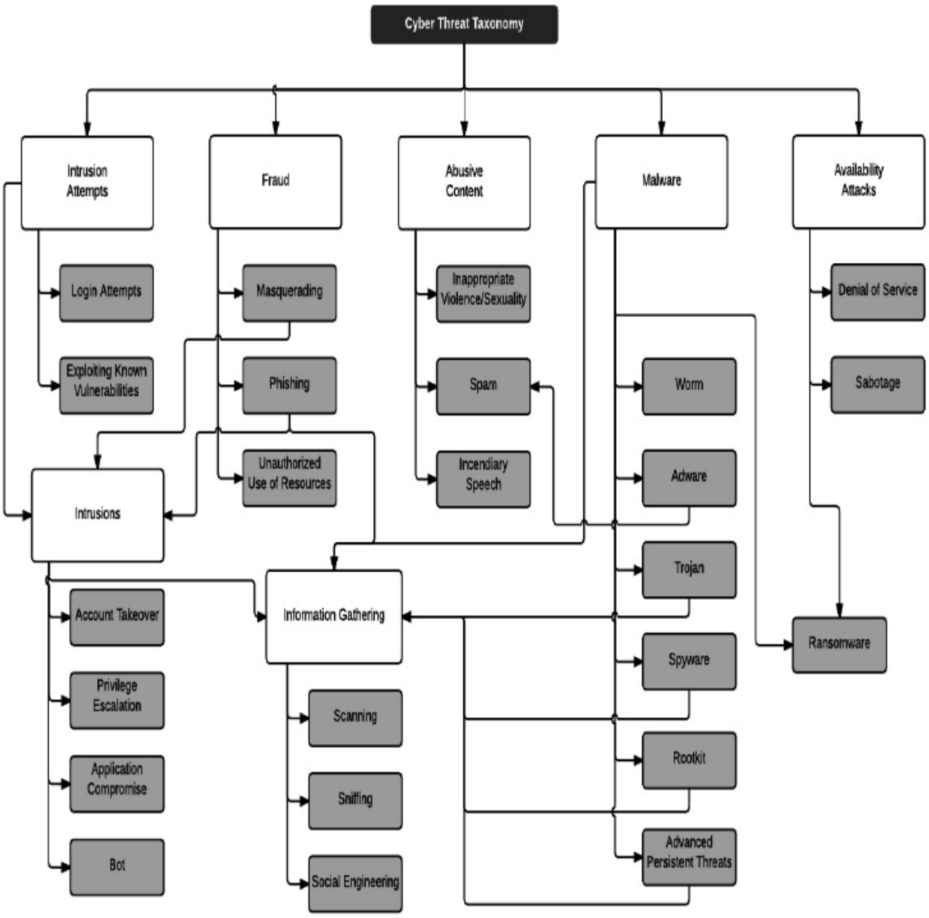

*Figure 1: CyberSecurity Threats Taxonomy*

Generally speaking, supervised learning models may make sense when there is a large amount of both legitimate and not legitimate cases with which to train the model. In other more challenging scenarios, it is quite difficult to find a representative set of positive cases that is sufficient to learn what positive events are like.

These are the cases, for example, of intrusion detection, breaches caused by zero-day attacks or new vulnerabilities, that are better approached by anomaly detection techniques.

### 3.1.3.1. Anomaly Detection

Anomaly detection has demonstrated to be a very effective way to early identify malicious activities covering a quite large variety of case, from denial of service attacks families (DoS) [21] to worms [22] and so on and so forth.

The anomaly detection module reveals traffic anomalies by comparing the run-time collected data with the learned normal behaviour based on the past traffic history and looking for significant changes in short-term behaviour (on the order of minutes to hours) that are inconsistent with the expected forecast. In state-of-the-art cases, the detection blocks typically treat the traffic as a collection of flows that need to be examined for significant changes in traffic pattern (e.g., volume, number of connections, specific counters, etc.)

The process behind predictive modelling can be summarized in the following points [23]:

- **Data identification**.
- **Data pre-processing**. It may involve data cleaning, conversion of categorical features to numerical values, data filtering, dealing with missing values, and feature engineering (creating new features).
- **Splitting the data into training and testing**. Usually the data are divided into training and testing sets either randomly or based on time, in cases where the temporal dimension of the data is important.
- **Algorithm selection**. There are many kinds of machine learning algorithms, differing mostly in how the decision boundary between the classes is decided and how complicated it may be. Examples include Regressions, decision trees, random forests, support vector machines, neural networks, etc.
- **Training & parameter tuning**. The model is trained on training data. Model parameters (hyperparameters) are tuned on training or on an additional validation data.
- **Model evaluation**. Model performance is evaluated on the test data set with respect to the metric established at the beginning of this process.

Going from a business problem to a predictive model is usually an iterative process that may involve experimentation at every step.

### 3.1.3.2. Threat Modelling

Threat modelling is the process of the systematic enumeration of threats to a system. Just claiming that a system is secure is not enough; it is necessary to explicitly define the attack vectors against which the envisioned security enablers are going to be tested.

There is no common security solution for all systems against all threats; each one has specific security requirements and possible threats, and it is therefore required that the enforced policies are specifically tailored for the particular system under analysis. When choosing the policies, the security engineer must have a holistic perspective of the service graph to be deployed with all the details such as the specific VFs, all the data flows, the access points, any privileged code and the defined trust boundaries. Threat modelling systemizes processes for identifying all the needed pieces of information that a systems security engineer needs for specifying and enforcing security policies.

Threat modelling produces an analytic schematic of the system that identifies all the possible threats, assesses them and rates them according to their possibility of occurrence and the relative impact they could have. This model helps to develop realistic and meaningful security policies that perfectly fit the specified requirements. It is therefore clear that threat modelling has a strong relationship with the definition of the security requirements of the system and the development of the security mechanisms (Figure 2).

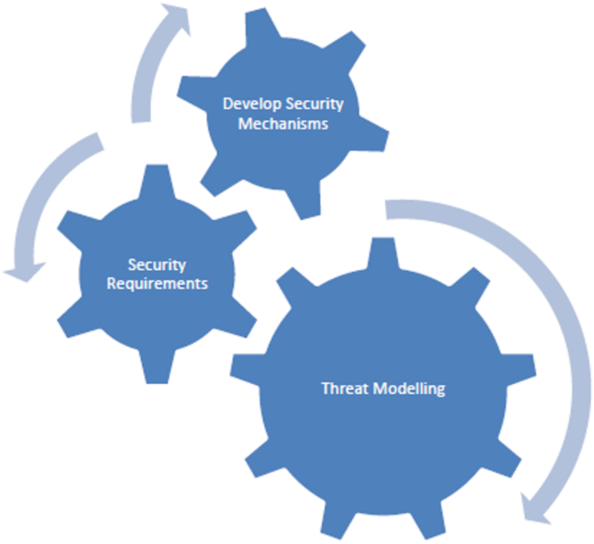

*Figure 2: Threat Modelling* ***Errore. L'origine riferimento non è stata trovata.***

There are several models that can be leveraged for describing any discovered threats and attacks against the deployed VFs and that are currently effectively used for visualizing threats both at the network and the host layer:

- **Attack Tree:** Attack trees can be used to model the possible attack paths that lead to specific attack goals. They use tree-like diagrams to represent the possible attack paths that are available in the system. More specifically, the root of each tree is a vulnerability identified to be a threat to the system, the leaves represent the attack surface and the rest of the nodes are subtasks that lead up to the realization of the attack. The modelling process with attack trees has multiple steps [24]:
  - Identify the threats that the system might have and set them as attack tree roots;
  - Decompose each threat to subtasks iteratively until the subtasks consist of basic functionalities which belong to the attack surface. Some tasks might be composed of all their subtasks, and some only of one, depending on this setting, there are logical AND and OR nodes to be used;
  - Each task is assigned a cost number. If the final attack has a smaller damage than its summed subtasks costs, then the attack is deemed not likely to occur. On the other hand, attacks with a low-cost implementation and highly damaging results, indicate that the attack must be mitigated as it is likely to occur.
- **Attack Library**: Attack libraries are collections of known attacks compiled into searchable databases. They are general purpose checklists, that aim to provide detailed information for attacks, in order to help threat modelers to understand each threat from the perspective of the attacker. Attack libraries can be used to check for known vulnerabilities against the implemented functionalities and cross them out from the list. Some major libraries are:
  - OWASP;
  - CWE;
  - CAPEC;
  - WASC Threat Classification
- **T-MAP** [25]: It is a quantitative threat modelling technique for assessing security risks by calculating the total severity weights of relevant attack paths. This method begins by identifying attack paths on a four-layer model. The layers consist of: the firewall, the

commercial of the shelf (COTS) systems, the IT infrastructure and the organization’s core values. These layers are described by 22 attributes which are derived from the Common Vulnerability Scoring System (CVSS). After that, weights are applied to each of those attributes depending on their severity, and each attack path can be evaluated based on those weights. The overall threat of the system can be quantified by summing the weights of each attack path, in order to provide a security evaluation of the whole system. Finally, the threat modeler, evaluates all the countermeasures for each threat, depending on their efficiency and cost. This way, optimal countermeasures can be chosen for the system and minimize both costs and possible damages. This method can provide comparative results against similar implementations in order to assess the threats against the system.

## 3.2. Networking

### 3.2.1. Network exposure capabilities beyond 5G

One of the new paradigms introduced by 5G and that offers an enormous capacity for innovation is its programmability through 5G Core. Specifically, exposing features and functions of the network through APIs to external entities (third parties), such as developers, to achieve a more secure and efficient access to network components while expanding the possibilities.

A key feature of 5GC is that the flexible nature of the architecture is service-based (SBA) [26]. This means that Network Functions (NFs) that compose 5GCore functionalities (AMF, PCF, AUSF, NSSF, etc) can communicate with each other and access their services if authorized, since service-based interfaces (SBIs) are exposed.

The continuous evolution of the 5GC led to the standardization by 3GPP of a flexible mechanism, the Network Exposure Function (NEF) [26], which its primary mission is to securely expose network data and capabilities to third parties. The NEF acts as an intermediate point between the southbound and northbound interfaces, which results in the creation of network APIs. These APIs are used, for example and among other things, for external AFs to modify the behavior of the network. Which, by the way, without proper monitoring and security features can result in a major security breach. In any case, there are clears benefits of network exposure through the NEF, for example, limiting the complexity of the underlying network, monetizing some network features, and a controlled access for external AFs.

A key part of the NEF is the NEF Northbound interface, which is a RESTful API in charge of many procedures between the NEF and an external AF. Among these procedures we can find the securitization of communication. The NEF Northbound APIs also have a service based-approach, which allows activities such as subscriptions to services or notifications to take place between the NEF and external entities using the APIs. NEF Northbound APIs are based on CAPIF (Common API Framework) [27], whose objective is to unify and standardize the use of exposed 5G capabilities. One of the most interesting features of CAPIF, related to the possible security breach mentioned above, is the authorization of API invokers. This is a way of ensuring that the entities using the network exposed functions are verified. Furthermore, the use of an unified framework such as CAPIF provides an abstraction layer which simplifies the heterogeneity of the network. Therefore, many applications do not need to be modified to use 5G capabilities.

The exposure of network functions and capabilities opens a very diverse range of scenarios both in the present and future, related, for example, to Network Slicing, ML, Edge Computing use cases, V2X, AR/VR and, in short, any scenario that may have exposed capabilities of a NF in 5GC.

Another 5G relevant function is Network Data Analytics Function (NWDAF) [28], used for collecting data (KPIs and information about different network domains), and provide analytics-based statistics to 5G core functions. While this is a powerful mechanism to introduce AI/ML and automation control-loops in the network management is also a risk, since a new point is exposed for various types of malicious attacks, and therefore the general security recommendations regarding 5G network elements should be applied here [29].

It is expected that this openness of 5G networks via the exposure functions, APIs for third party vertical applications and the general trend of using AI in network management will continue expanding more as part of 6G. Thus, it is critical to address security topics in the design of 6G architecture and not be and afterthought, considering potential vulnerabilities associated to the exposure APIs, network data and the applicability the AI to network automation.

### 3.2.2. Energy Efficiency

The full deployment of 6G networks will necessitate the development and deployment of energy efficient ML algorithms that can support the latency vision of 6G with minimum hardware complexity and energy consumption. To this end, federated learning (FL) has been proposed as an efficient way to deal with the training of large and diverse datasets and at the same time mitigate potential security and privacy concerns [30], [31]. In this context, a set of independent nodes is assumed that performs local training based on the available data set to extract an ML model. The parameters of this model (i.e., weights of trained neural networks) are send to a centralized location where model aggregation and update is performed periodically. Hence, instead of transmitting the entire data sets to the central processing node, only the corresponding weights are sent. Consequently, in FL training mode, the overall computational burden is divided among the participating nodes, thus reducing energy footprint. Moreover, in conventional centralized training, data sources can be sparsely distributed, which in turn require an increased amount of transmission power for the proper training of the ML model with all available samples. On the contrary, in the FL case, since only the corresponding weights are transmitted, overall transmission power can be significantly reduced.

Another key concept towards energy efficient 6G networks is the optimization of network infrastructure. This involves the use of energy-efficient hardware, such as base stations and routers, as well as the deployment of energy-efficient networking protocols, such as Software-Defined Networking (SDN) and Network Function Virtualization (NFV) [31], already adopted in the 5G initial deployments. To this end, in SDN the elements in the control and forward plane can be separated, thus providing a more robust mechanism for packet delivery which reduces energy consumption. This is made feasible in conjunction with the NFV technology since software and hardware elements can be decoupled [32]. Thus, the concept of network on demand is supported, and all related functionalities are moved closer to the end users.

In the same context, distributed solutions such as cloud-RAN (C-RAN) can leverage optimum network deployment, since due to their dynamic nature propagation distances and overall transmission power can be reduced [33]. This type of cloud computing environment operates on open hardware and network interface cards that dynamically handle fiber links and interconnections within the station. C-RAN is more cost- and footprint-effective due to less hardware, which in turn leads to lower heating, cooling and power requirements.

Small cells have also been recently proposed as an efficient way to reduce transmission power, especially in dense 5G networks, due to their ability to provide an extension from the core network into densely urban areas. Small cells can be a quite attractive solution in highly populated scenarios, such as crowded stadiums or large venues, shopping malls, and industrial buildings [34].

Finally, Symbiotic radio (SRad) technology is a promising solution that enables cooperative resource-sharing among radio systems through symbiotic relationships. By fostering mutualistic and competitive resource sharing, SRad technology enables the achievement of both common and individual objectives among the different systems [35]. In this context, one radio system is passive and the other is active and both share the same resources, where the passive radio system depends on the resources of the active radio system e.g., passive IoT and WiFi. Hence, a more collaborative approach can be formulated in terms of transceiver procedures, which reduces transmission power and leverages energy efficient architectures.

### 3.2.3. Digital Twin design

Digital twin networks are a type of cyber-physical system that simulate and model physical assets or processes in real time. These systems are becoming increasingly popular in industrial sectors such as manufacturing, healthcare, and transportation, as they can help optimize processes, reduce costs, and improve overall efficiency. One of the most significant trends in digital twin networks is the integration of machine learning and artificial intelligence. By incorporating machine learning algorithms, digital twin networks can become more intelligent and autonomous, enabling them to make decisions and predictions based on real-time data.

Digital twins for computer networks involve creating a virtual model of a network infrastructure that can be used for monitoring, troubleshooting, and optimizing network performance. These digital twins are typically created by collecting data from the network in real-time and using machine learning algorithms to build a predictive model of network behaviour.

One of the key advantages of using digital twins for network management is their ability to provide real-time visibility into network performance. By monitoring network traffic and behaviour in real-time, digital twins can help network operators identify and diagnose issues more quickly, allowing them to respond and resolve problems faster than they could with traditional methods.

Another advantage of digital twins for network management is their ability to simulate and predict network behaviour. By building a predictive model of network behaviour, digital twins can help network operators optimize network performance, identify potential issues before they occur, and even simulate the impact of network changes before they are implemented.

Digital twins can also be used to improve network security. By modeling network behaviour and identifying anomalies in real-time, digital twins can help network operators detect and respond to cyber threats more quickly and effectively than traditional security measures.

Key challenges associated with digital twins for computer networks include the complexity of network infrastructure and the need for large amounts of data to build an accurate predictive model. Additionally, digital twins must be updated and maintained to reflect changes in the network infrastructure or configuration, which can be a time-consuming and resource-intensive process.

As proposed by Autodesk[1], there are five levels of digital twin, where each layer requires a greater degree of maturity and digital transformation, but also includes increased value. The five levels are:

[1] https://f.hubspotusercontent40.net/hubfs/4545544/AEC%20Demystifying%20Digital%20Twin%20(EN).pdf

- Descriptive Twin, which provides a live, editable version of design and construction data of the physical twin;
- Informative Twin, which provides additional operational modelling and sensory data;
- Predictive Twin, which leverages operational data for insights;
- Comprehensive Twin, which provides simulations for future “what-if” scenarios;
- Autonomous Twin, with the ability to learn and act on behalf of the users.

Recently, the usage of Digital Twins (DTs) was extended to mobile networks, and especially 5G and 6G networks. The main objective of such works is to enable the integration of AI/ML loops within the management process in order to increase autonomous behavior and reaction to faults, as well as to improve and optimize the network performance.

At the moment, the topic is not yet completely mature in the scientific literature. However, some interesting and relevant works include:

- In [36], the authors describe at high level the usage of digital twin technology to integrate AI solutions within the network life cycle. The paper describes how DTs enable to learn, optimize and test the network during the design and development phase, how provide collective intelligence during operation, and how to support knowledge transfer during the expansion and extension of the infrastructure.
- In [37], the authors provide a holistic view on the usage of DTs in 6G, as they explore the applicability of the DT technology in the context of 6G communication systems by viewing it as a tool to make research, development, operation, and optimization of the next-generation communication systems highly efficient.
- [38], instead, focuses on the usage of DT technology in security scenarios. The paper presents the usage of a simplified DT (i.e. a cyber range) to augment cyber range environments with ML tools. The work focuses on how this scenario might work and the way in which it might be used to train new experts.

### 3.2.4. Physical-Layer in 6G Networks

The 6G mobile communication networks should satisfy strict requirements related to reliability, latency, and security. Simultaneously, they should also provide a significant improvement in coverage, data transfer rates, user experience, and network capacity. As a result, the key performance indicators (KPIs) that will be adopted are expected to be 10 to 100 times better than those employed in 5G. As far as the physical layer is concerned, novel communication techniques are planned to be employed, related to the following scientific pillars i) higher frequency bands, ii) intelligent reflecting surfaces, iii) hybrid massive multiple-input multiple output (MIMO) communication architecture [39]. As far as the first pillar is concerned, currently an important part of the research in the physical layer focuses on i) Millimeter-wave technologies; ii) THz technologies; and iii) Free space optics (FSO) [2]. The new spectrum that will be used is expected to considerably increase the available bandwidth in order to support the evolving and capacity hungry 6G use cases that include augmented and virtual reality, unmanned mobility, and e-health [40].

Regarding the investigation for the smart radio environment, the use of reconfigurable intelligent surfaces (RIS) shows great potential for wireless communication networks. More specifically, RISs are nearly passive surfaces made up of electronically adjustable reflecting elements, which offer a way to reconfigure the incident signals, in order to create virtual line-of-sight propagation conditions for wireless transmission [41]. From a practical perspective, various approaches have been proposed for controlling the electromagnetic wave characteristics including dynamic reflecting arrays, tunable metasurfaces, and liquid crystal surfaces [42]. By optimizing the electromagnetic wave propagation environment, RIS can enhance both the spectrum and energy efficiency of wireless networks. In [43], a new communication technology has been proposed that combines the index modulation principle with RIS for further improving the spectral efficiency. In [44], a RIS-assisted mmWave communication system has been investigated and a new channel model has been introduced, which can be applied in various communication scenarios.

The realization of everywhere and every time connectivity in 6G era cannot be relied only in the terrestrial cell-based architecture or the traditional MIMO technology. For example, in various scenarios, e.g., marine, rural or in crowdy events, other solutions should be adopted to support the 6G requirements that are based on non-terrestrial solutions or cell-free massive MIMO techniques [2]. Towards this objective large scale satellite constellations in low Earth orbits can offer increased data rates in relatively low latency. Moreover, high altitude platforms or un-manned aerial vehicles could be also used to provide telecommunication services in a cost-efficient manner, offering near line-of-sight propagation conditions and dynamic adoption to the traffic requirements. Moreover, a new air interface technology utilizing non-orthogonal multiple access (NOMA) has also been proposed to enable massive connectivity in the 6G era.

Moreover, 6G communication networks requirements for hyper security in the vast number of devices that will be always connected impose several challenges. To this aim, the physical layer security (PLS) is expected to provide an additional and quite effective degree of freedom that will address the upcoming challenges for supporting higher bandwidths, higher carriers, lower latencies, reduced energy consumption, all with secure manner [45]. In this context, various contributions have been reported that deal with PLS in different communication scenarios. The authors of [46] proposed a maximization framework for the secrecy rate for wiretap channels using RIS-assisted multiple-input single-output (MISO) technique. Moreover, the authors of [47] improved the PLS for RIS-assisted non-orthogonal multiple access (NOMA) in 6G networks. Focusing on dynamic scenarios, the authors in [48] analysed the secrecy capacity of RIS for vehicle-to-infrastructure (V2I) communications. Additionally, since artificial intelligence (AI) is recognized as one of the enablers of 6G communication networks, various contributions have used AI tools to improve the performance of the systems under investigation. For example, in [49], AI has been utilized to create intelligent and robust security solutions in 6G networks. In [50], an adaptive security specification method has been proposed, which was based on AI, in a 6G Internet of Things communication scenario. Moreover, THz and mmWave frequency ranges have been employed to provide security protection for different services. In [51], a deep learning method has been proposed for maximizing the average secrecy rate of a UAV communication system with partial channel state information.

## 3.3. Artificial Intelligence

### 3.3.1. AI-enabled solutions for security enhancement in 6G and threat mitigation

AI play a critical role in 6G, not only in the design and optimization of protocols and operations, but also in the design of early detection of threats and anomalies. AI benefit 6G security systems, but the alliance between 6G and AI is a double edge-sword as it can also become a target of attacks.

Unlike 5G networks, where security solutions across all devices and base stations are configured with universal settings for certain types of attacks, it is apparent that such an approach cannot be applied in 6G networks. Intrusion Detection Systems (IDS) have been extensively used, but it is demonstrated that they fail in the detection of complex attacks. Cybersecurity attacks in 6G networks are dynamic, polymorphic and sophisticated, using previously unseen custom code, able to communicate with external command and control entities to update their functionality. To this end, a smart support system is required for the prediction of attacks, detection of threats and the definition of proactive actions, previous to mitigation strategies. This will require the evaluation of the impact of the attack, the criticality and resilience of the infrastructure compromised and the cost of the proactive actions and effective mitigation.

[52] reviews security and privacy issues of 6G networks in the physical, connection, and service layers. It identifies new threat vectors, different from 5G, such as threats in the physical layer, and security issues in distributed AI. In [53], the authors have proposed an optimisation framework to address the identified challenges in 6G networks. The proposed framework optimizes security scheme selection and configurations to balance the security-energy trade-off in various scenarios. In [54], the authors analyse various potential new threats caused by the introduction of new technologies related to the usage of open-source tools and frameworks for 6G network deployment and present possible mitigation strategies to address these threats.

AI started to be used by security solutions [55] to overcome their limitations in the detection of complex and zero-day attacks. Initially, signature-based and anomaly-based detection systems made extensive use of classical ML techniques, however they lack automatic feature engineering, they have a low detection rate, and they are not efficient in detecting small variants of existing attacks. Consequently, DL techniques were adopted forced by the increasing complexity of hacking incidents, zero-day attacks, and unknown malware. DL based security solutions have been successfully developed on the infrastructure level (intrusion and anomaly detection), software level (malware, virus and botnet detection) and privacy level (personal information).

Nowadays, 5G networks make use of ML techniques to achieve dynamic and robust security mechanisms [56]. Moreover, distributed learning [57], Deep Reinforcement Learning (DRL) [58], and Federated Learning (FL) [59] models have proven their ability in the detection of complex and zero-day attacks in distributed environments.

According to the existing state-of the-art, 6G networks are more prone to different types of cyberattacks and security threats. Either the network is real or it is based on digital twin (DT) [60]. To detect the cyberattacks, machine learning and deep learning methods based intrusion detection tools are used [61].

Several surveys [62], [63] in the literature have analysed the usage of AI by security applications proving that they are suitable for 6G security enhancement. In the physical layer, AI can improve the performance of the detection engines using DRL [64] to enhance randomness in physical layer phase-modulated key generation and to enhance physical layer authentication [65].

At the network layer AI is considered to enhance security solutions performance in the prediction of network attacks [66], filtering malicious traffic and making intelligent recommendation for network changes. Finally, in the service layer, AI is a favoured technique in different aspects, such as access behaviour modelling [67], or in biometric authentication [68].

DL methods can provide big data security support to the 6G end-to-end system which includes cross layer optimization such as optimizing the channel coding, synchronization and estimations [69]. Distributed solutions are more focused on the edge and end to end solution while securing the 6G networks. The edge DL solutions are not capable of handling all types of attacks. To handle this issue meta learning approach is proposed which adaptively change the ML model running on a device to improve performance and accuracy of the model. In article [70], authors have utilized the meta-learning algorithm to identify the Wi-Fi impersonation attack. In cyber physical systems (CPS), intrusion detection systems are developed by using Federated Deep Learning (FDL) algorithm [71]. The proposed models support the multiple industrial CPS while preserving the privacy of the system. Both meta-learning and federated learning (FL) are used to detect cyberattacks in the 6G networks but these approaches are not mature enough to guarantee privacy of the physical device [72].

In [73], the authors have identified the 23 different attacks that are experienced by the integration of IT (Information Technology), OT (Operational Technology) and IIT (Industrial Internet of Things) on the surface of the 6G network. They have used Programmable logic controller with factor I/O to simulate the different types of attack. Then they have developed a solution based on Deep Learning (DL) models, RBM and RNN and incorporating it with LSTM architecture to detect malware and ransomware families of attacks.

Another important issue of current AI solutions is the usage of centralized data which poses serious privacy issues [74], while it is not aligned with the distributed architecture of the 6G networks. Moreover, ML and AI-based optimisation approaches can be used to improve time-series and statistics-based methods to operate beyond non-normal and train the system generating attacks. For instance, in [75] the use of Generative Adversarial Networks (GANs) is explored to simulate intrusions and malware for improving its detection, and to fuel defense against different attack methods. While, in [76] Transfer Learning (TL) is used to improve the detection of zero-day attacks.

### 3.3.2. Intent-based Networking

An intent is defined as a set of operational goals (that a network is supposed to meet) and outcomes (that a network is supposed to deliver) defined in a declarative manner without specifying how to achieve or implement them.

Intents define goals and outcomes in a purely declarative way, stating what is to be accomplished, not how. Thus, intents apply several important concepts simultaneously.

- **Provides data abstraction:** users don't have to worry about low-level device configuration and nerdy controls.
- **Provides a functional abstraction of specific management and control logic:** users don't even have to worry about how to achieve a specific intent. A desired outcome is specified, and IBS automatically determines the course of action on how to achieve that outcome.

In the context of autonomic networks, the translation of intent into device-specific rules and actions ideally should be performed by the network itself. This decentralized approach would rely on distributed algorithms and local device abstractions, allowing intent to be automatically disseminated across all network devices. However, in certain cases, some level of centralization is necessary. For example, a conceptual point of interaction with the network may be required for users, which can act as the operational front end and interact with other systems in the network to fulfil the desired intent. Additionally, a centralized approach may be beneficial when global knowledge of the network state is necessary for achieving certain intents.

In many situations, specialized functions or dedicated systems are needed to provide intent functionality. These functions may handle the translation of specific types of intent into corresponding actions and algorithms, and they can be implemented in a distributed manner to avoid single points of failure. Regardless of the implementation, an Intent-Based Networking (IBN) is a network that can be managed using intent, recognizing and adapting itself to user intent to achieve the intended outcomes without requiring detailed technical instructions. Similarly, an Intent-Based System (IBS) serves as the point of interaction with users and implements the necessary functionality to achieve intended outcomes, interacting with the network as required.

Intent-based networking (IBN) [77] is an approach to network management and automation that focuses on aligning network behaviour with business intent. It aims to simplify network operations, enhance agility, and improve network efficiency by abstracting network configuration and management from low-level technical details.

Intent-Based Networking involves a wide variety of functions that can be roughly divided into two categories: Intent fulfilment and Intent Assurance.

Intent fulfilment involves functions that ingest, translate, and orchestrate intent across the network. Intent assurance involves functions that monitor, assess, and report on the compliance of the network behaviour with the intent. Both fulfilment and assurance functions may involve learning and optimization capabilities to improve the effectiveness and efficiency of intent realization.

# 4. HORSE Use Cases

## 4.1. HORSE Use Case 1 - Secure Smart LRT Systems (SS-LRT)

### 4.1.1. Use Case Description

Light Rail Transit (LRT) or Metro Operation involves the management and orchestration, with high availability, of several systems, applications and end-to-end services, supported by equipment that typically are deployed on tram stops, trams and in the Command Centre (OCC). Usually, these Command Centres is deployed in private networks for security reasons and are located in the Operator premises, for security and for latency reasons.

Therefore, based on the Dublin/LUAS LRT scenario, architecture and functionalities, this use case intends to compare the performance of the traditional or common system with the one supported by 6G capabilities, achieved by the innovative HORSE solutions. Moreover, it is expected to overcome some known limitations such as congestions, threats and vulnerabilities related to cybersecurity or even disruptions based in the new network capabilities.

The Metro Operation scenarios are expected to benefit from new paradigms related to communications, disaster recovery, security, and resilience. The geographically distributed operation (even supported by cloud solutions) will pose a significant impact on the overall availability and the decision support.

### 4.1.2. Problem Statement and adaptation to the HORSE Infrastructure

Considering Metro infrastructures and Metro Operations, the Cyber Security is one of the biggest problem statements. According to ENISA Threat Landscape: Transportation Sector [78], based on real cases during 2022, the most frequent attacks are:

- Ransomware attacks (38%)
- Data related threats (30%)
- Malware (17%)
- DoS/DDoS (16%)

- Phishing/Spear phishing attacks (10%)
- Supply-chain attacks (10%)

For the OT – Operational Technology environments the Cyber Security vulnerabilities are very common and easy to be exploit, by attackers, because generally the network has several legacy equipment and has a long spatial distribution. The consequences can lead to disruptions in E2E services with impact in the operation, whit impact in the passengers, damages in the image of the infrastructure owner or in the Metro Operator and significant losses. Typically, the network is distributed via TETRA, WIFI, Ethernet cable and fibre optics along a very wide area that is limited for the terminal stations of the Metro System.

For each kind of attack enumerated and quantified in the ENISA report, follows a brief description of how the use case scenario could be affected by the threat and what data of the organization is vulnerable:

- **Ransomware attacks:** Ransomware attacks in the use case scenario could be done via installation of malware on USB pens on workstations or servers. The malware can be undetected for days/weeks while trying to lateral move for other machines in the network in order to affected them with the malware. Once the malware is installed in all machines available in the network, it can be triggered in a more convenient time where the network traffic and processing volume is lower to encrypt all the files more silently in the machines without being detected. When all the files are encrypted the ransom process could be deployed.
- **Data related threats:** Threats against data in the use case scenario can be deployed by exploitation of vulnerabilities of OS where DB is running. Because and assuming that VLAN are well configured, it is not possible to suffer from external attacks from the internet. But internal attacks can occur and OS without updates are vulnerable to exfiltration of data from DB and this data integrity is important for instance for payment evidence of voyages timetable performance in relation to timetable or data like CCTV images or even PIDS or audio announcements deletion or modification.
- **Malware:** Malware attacks in the use case scenario could happen via installation of malware in an internal machine without anti-virus active and at the end the malware could end in a ransomware attack. The malware can be installed via an infected USB pen or by lateral movement between infected machine in the network.
- **DoS/DDoS:** DoS/DDoS attacks in the use case scenario could be deployed via internal network machines or external internet attacks that can block access of use case services to the internet. Another occurrence can be the TETRA or WIFI signal jamming or traffic exhausting the service.
- **Phishing/Spear phishing attacks:** Phishing attacks in the use case scenario could be affected by social engineering attacks that steal login information on systems from operators and then attackers have privilege access to confidential information in the internal machines. EFACEC machines could be affected by other machines in the same network that were attacked by phishing attacks.
- **Supply-chain attacks:** Supply-chain attacks in the use case scenario could be triggered in equipment without updates connected in the network like, switches, routers, CCTV, PID, etc. and are vulnerable to internal or external attacks.

Some other issue must be considered, since the use of the HORSE network is based in open communications (wireless) instead of close communications One example of vulnerability is related to signal jamming.

### 4.1.3. Demonstration: Usage Scenarios

The Use Case was structured in several scenarios taking in consideration different functionalities, different requirements and by inherency different validation processes. The use case and the proposed scenarios are based in two important Metro operations, such as Public Information systems and Security Systems involving the following functionalities:

- Digital signage and public addressing (stations and onboard), that can be affected by ransomware, malware, phishing, data related threats, DoS/DDoS and supply-chain attacks.
- real-time security CCTV systems (stations and onboard) to help dispatchers at operational control centres, drivers, and security agents activities at station's platforms, that can be affected by ransomware, malware, phishing, data related threats, DoS/DDoS and supply-chain attacks.
- Passenger's video help point assistance (stations and onboard) that can be affected by data related threats, DoS/DDoS and supply-chain attacks.
- Security and maintenance agents' awareness (stations and track) that can be affected by data related threats, DoS/DDoS and supply-chain attacks.

The HORSE solution must minimize and help to mitigate the cybersecurity threats, must assure an efficient level of slice orchestration in order to assure the proper level of security, latency, avoiding congestions, independence of E2E services and disruptions.

All these scenarios, require that the selected equipment and devices support 5G/6G CPE to assure the communication to the HORSE network.

Besides these UC demonstrated scenarios, that will support the validation of the HORSE solution some other advanced scenarios will be considered as case studies in order to validate the future trends of 5G/6G to support Metro/Rail Operations. New radio networks that can be used instead of Tetra networks or even GSMR networks leveraging FRCMS or 5G/6G to be the future of integrated and convergence networks to support Metro/Rail Operations.
These new solutions can overcome the following restrictions:

- Tetra networks to communicate between the OCC and the Tram – Only support voice and short data messages. Since there are bandwidth limitations typically it is only used to transmit localization information. A ORSE solution can be exploited to transmit some other kind of data (operational statistics, service/timetable information, for instance) and voice and video messages at the same time.
- Private radio networks to communicate between the track and the tram, typically for localization and signalling purposes. A HORSE solution can also be considered to support this kind of communications.

Therefore, adopting a HORSE solution could lead, in the future. to a unique network in to support Metro/Rail operations, improving the convergence, integration, the costs of maintenance and even the CAPEX and the OPEX of a Metro/Rail solution also assuring the capability to support new type of services.

#### *4.1.3.1. LRTUC1 - Public information systems*

This scenario involves the communication between the OCC (Operational Command Centre), deployed at Metro premises or in the Cloud, and specific displays (PID-Public Information Devices) located at tram stops or at the trains (specific onboard displays at passenger's cabin). Typically, these operations are necessary when the Operator intends to broadcast free text or pre-recorded messages to the PIDs or when it is intended to broadcast the forecast information (arrival and departure time information) to the tram stops. Additionally, if the trams are equipped with special devices, it is possible to send messages containing advertising or even video. The HORSE platform must be able to manage requirements related to multicast, security and performance (low latency) to assure that the messages can reach the devices with minimum delays and the same time.

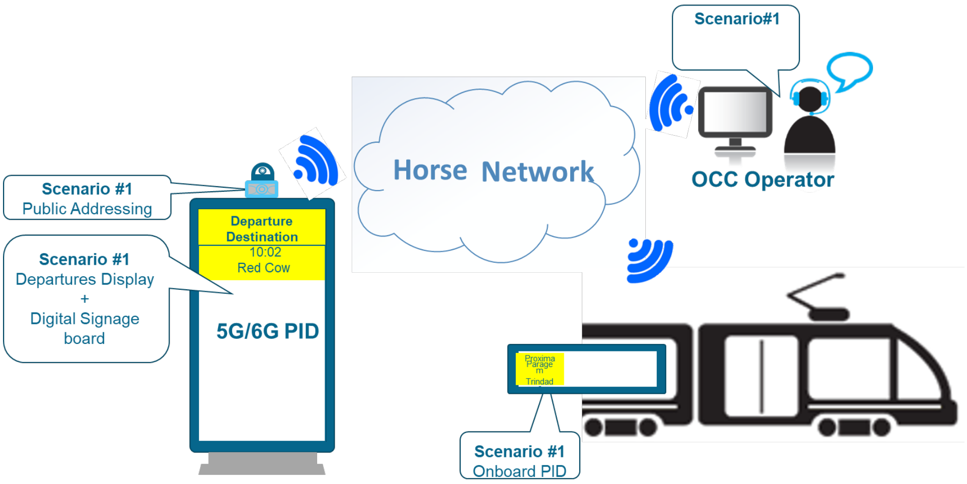

*Figure 3: Public Information System, supported by the HORSE Platform*

### *4.1.3.2. LRTUC2 - Real time security video transmission*

This scenario is based on the video capabilities to reinforce the security of the Metro areas both at tram stops, including the track, and at the onboard tram. Therefore, the video solution allows the fixed communication between the tram stops or even the track and the OCC, allowing the operators to visualize, in real time, images of the requested areas. Moreover, this scenario allows the operators to visualize video images from the train (mobile communications) both from the passenger's cabin area and from cameras installed in the driver's cabin. The cameras installed in specific areas, allowing the visualization of the tram stops and the track can be transmitted to the driver in order to detect the presence of people or obstacles, in real-time. This scenario also requires a reliable horse platform able to address multicast, security, and low latency capabilities.

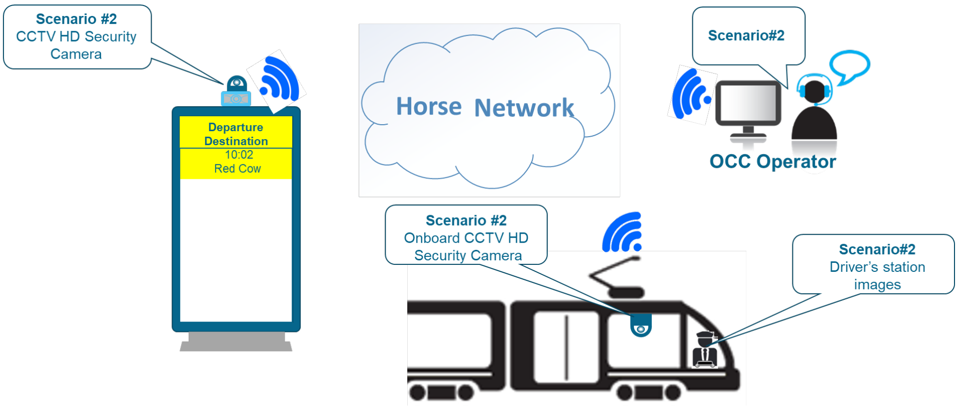

*Figure 4: LTUC2 scenario supported by the Horse platform*

### *4.1.3.3. LRTUC3 - Assistance to the passengers*

The main goal of this scenario is to assist passengers located in tram stops or inside the train, based on dedicated devices (help points) supporting live conference, allowing the communications between the passengers and the OCC. Therefore, if a passenger needs assistance (regular information or in an emergency situation), they can use the help points to communicate with the OCC. A secure communication channel is established in a point-to-point fashion. The passenger is able to report the emergency to the agent at the OCC, and at the same time, the agent at the OCC is able to assess the situation remotely through video and audio. If further assistance is needed, the agent at the OCC can check the train's location or the tram stop location and send specialized personnel to assist or resolve the incident.

This scenario requires a reliable HORSE platform able to support video and voice communications, to assure secure communications and low latency.

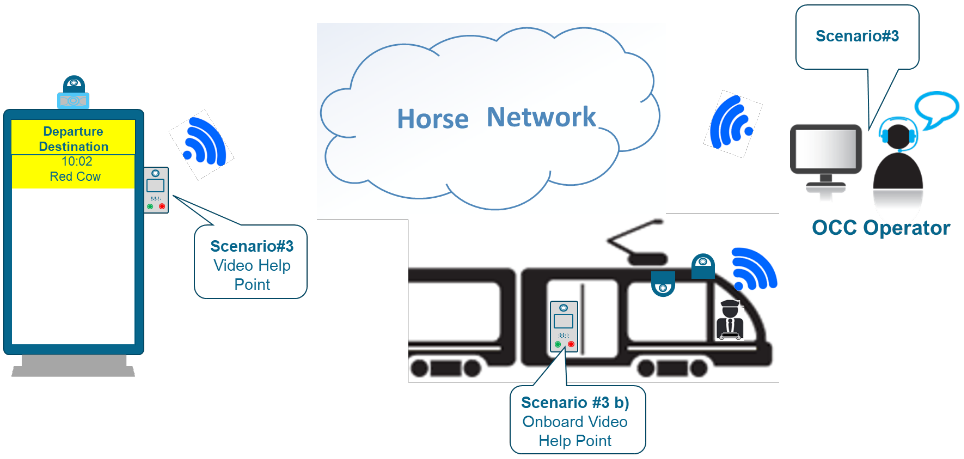

*Figure 5: LTUC3 scenario supported by the Horse platform*

### 4.1.3.4. LRTUC4 - Support operation: security and maintenance agents' awareness

This fourth scenario intends to reinforce the security and maintenance capabilities of authorized people that work at train stops.

Regarding the security, the main goal is to allow the security agents at tram stops to visualize, in real time, video images from approaching trams or video images from the track. For this purpose, the security agent uses a tablet/mobile device for receiving the video images.

Concerning the maintenance, the agent, using a similar device will be able to receive alarms or operational critical events related to the track or the technical rooms near the tram stop.

This scenario requires a reliable horse platform able to support video and data messages transmission, to assure secure communications and low latency.

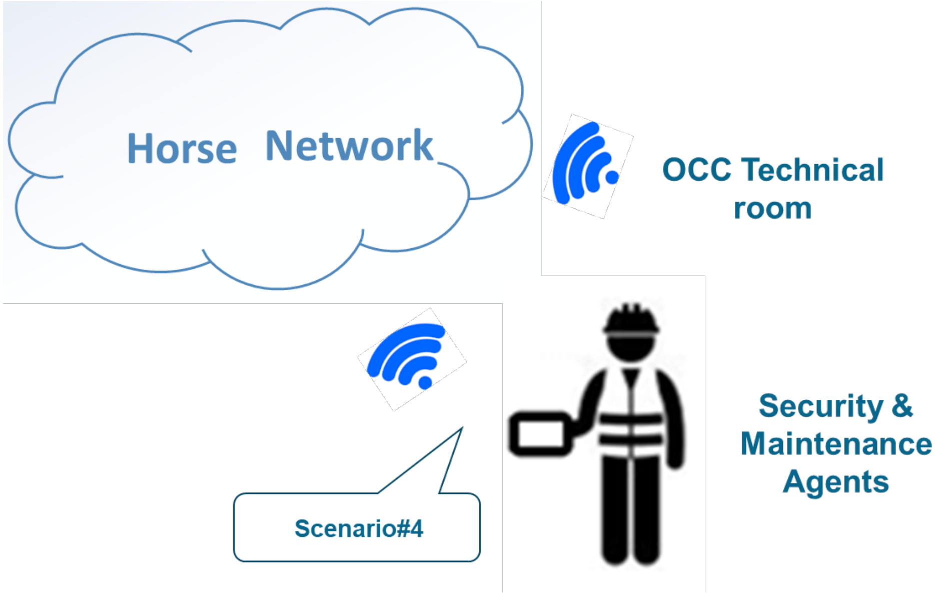

Figure 6: LTUC4 scenario supported by the Horse platform

## 4.2. HORSE Use Case 2 - Remote Rendering to Power XR Industrial ($R^2$2XRI)

### 4.2.1. Use Case Description

The second use case of this HORSE project (Remote Rendering to Power XR Industrial) will focus on remote application rendering in the Extended Reality (XR) industrial sector. The use of XR technologies has become commonplace in many industrial verticals, who aim to leverage the various beneficial properties of XR to enhance workflow processes. Among such enhancements, some of the key benefits are immersive training measures, remote support, and product design. An ongoing challenge in industrial XR is in the accommodation of the continued growth of these technologies and their associated business demands with appropriate network infrastructure and connectives.

This use case will utilize the HORSE platform to assist in meeting these network needs. The XR application *Augmented Reality Engineering Space (AR 3S)* by partner HOLO provides end-users with the ability to engage in multi-user virtual fast-prototyping. Fast-prototyping involves the collaborative display and interaction with a computer-aided-design (CAD) file virtually. Incorporated into AR 3S is a remote rendering and application streaming functionality, imparted by the Interactive Streaming for Augmented Reality (ISAR) SDK. Remote rendering and streaming provide high resolution and quality XR experiences by bypassing device processing limitations. This remote rendering ability is critically dependent on low latency, high throughput, and secure network connectivity. The various functionalities and attributes of the AR 3S XR application will be enhanced in this use case by leveraging the HORSE platform, which will aid in validating the components of the platform itself.

### 4.2.2. Problem Statement and adaptation to the HORSE Infrastructure

As mentioned above, XR technologies in industrial use cases are critically dependent on having a robust, resilient, and secure network infrastructures. This is particularly true for remote rendering and application streaming. During such a process, stable connections that provide low-latency and high-throughput data flow are essential to ensure the reciprocal exchange of rendered content and XR device sensor data is uninhibited. As innovations in the field of XR continue to blend the virtual and real world more seamlessly, largely aided by innovations in XR technologies and integration with Internet of Things (IoT) components, existing network infrastructures are challenged to accommodate increased data transmission and infrastructure needs. Additionally, as industrial sectors continue to further integrate XR technologies into daily operations, the density of high-volume data transmission will grow. These needs must be met with powerful and advanced network infrastructures, particularly in the context of 6G. Such connectivity solutions must also accommodate high levels of security and protection against the threat of outside attacks. Network security is critical in XR, as proprietary information in the industrial sector must be protected. The HORSE platform, in its mission to create an advanced infrastructure of novel services in the framework of 6G will assist in addressing these challenges in the field of industrial XR.

Within the context of XR remote rendering and application streaming, data exchange between a remote server (e.g., physical workstation, virtual machine) and the XR device (e.g., HoloLens 2 AR glasses) represents the most important information workflow. The specific data types and workflows of this information exchange are described in the following usage scenarios, which outlines the needs for the HORSE platform with respect to remote rendering in a local network. A key feature of these workflows is in the transmission of rendered XR content to the XR device, and sensor data which is sent to the remote XR processing server. The following sections will describe this in more detail.

### 4.2.3. Demonstration: Usage Scenarios

In this use case, several specific scenarios will be utilized. Each scenario represents a different aspect of the XR fast-prototyping through remote rendering experience with emphasis on AR 3S functionalities and their network connectivity needs which will be bolstered by the HORSE platform. Four scenarios will be assessed:

- Rendering of XR in a local network
- Fast-prototyping sessions in multi-player mode
- Multi-user experience
- Industrial Metaverse and XR devices

#### *4.2.3.1. XRUC1 - Rendering of XR in a local network*

Remote application rendering is a two-component solution that consists of a Server Application and a Client Application. The main application, containing the app logic, the interactable user interface elements, etc. are running on the Server (which can be hosted on a physical workstation such a laptop, or on a VM or on the Cloud). In addition, a lightweight Client Application is running on the XR end device glasses. These two components are connected via the WebRTC protocol and initiate a connection through a signalling event. Below, a brief explanation of how the communication functions:

A user wishes to visualize a CAD model using the XR device, e.g., HoloLens 2. The user uploads the XR application to a remote host device and starts an instance of the XR application. Following, the user connects the XR device to the wireless network with access to the remote computer where the XR application is running. In the next step, the user starts the client application installed on the XR device of choice. As the user moves, the XR device collects sensor input data (e.g., data about the movement of the user) and sends this data to the remote computer (I.e., the XR Server Application). The server, in turn, takes the data sent by the XR client device and renders a new view of the scene and sends this back to the XR client device. The application stream from the XR server is continuously sent to the XR client. This reciprocal data exchange allows the visualization of, and interaction with, holographic CAD files by users.

This reciprocal data exchange repeats while the user is interacting with the 3D model. At the end of the visualization, the user closes the client application in the XR device, disconnects from the network and finalizes the XR application running on the remote computer. The connectivity between server and client is critically important throughout the entire remote rendering data exchange process. During this remote rendering process, a robust and resilient network connection is required. This usage scenario will therefore directly exploit the 6G infrastructure and connectivity provided by the HORSE platform.

#### 4.2.3.2. XRUC2 - Fast Prototyping Sessions in Multi-Player Mode

Fast prototyping requires an application that can load, view, and manipulate 3D CAD models through smart glass enabled input. To allow collaborative product fast prototyping sessions, engineers have to be able to manipulate and compare entire 3D CAD models as well as specific parts in near real time. Importantly, to facilitate such a collaborative session, several end-users must be able to visualize the same model at the same time regardless of their physical location. During multi-player scenarios, several designers will visualize and interact with the same CAD file together. Multi-player collaboration of this nature also permits communication exchange, such as in the form of verbal discussion or leaving virtual notes in the virtual space. The shared space with which these users interact in must also be saveable, so the work can continue at a later stage by different stakeholders if necessary. Information exchange with multi-player sessions is dependent on network infrastructure which can handle the combined interaction data flow between participants in the shared sessions. A robust network connectivity is crucial for these interactions, particularly as the number of players within the shared session increases.

This use case scenario will utilize a multi-player session with multiple participants who will collaboratively interact with a visualized CAD file during a fast-prototyping session. This scenario will leverage the HORSE platform to ensure that the necessary network requirements are met and upheld during demanding multi-player sessions.

#### 4.2.3.3. XRUC3 - Multi-User experience

An additional scenario will focus on evaluating the permissive scope of the network during a multi-user environment. Unlike the multi-player session described above, a multi-user environment does not include several participants working collaboratively together on one CAD file in a single session. Instead, users will work on their own independent CAD files in their own respective sessions without interaction with others. These independent users will have their own application instances but will share the same network environment, such as when in the same building location. This type of use case scenario is common in XR environments where multiple end-users will work on their own task individually. The HORSE platform here will aid in ensuring that connectivity requirements are permissive during the simultaneous demands of individual AR 3S applications running on the same network. This use case scenario will consist of at least 3 end-users and their associated devices/application instances running at the same time.

#### 4.2.3.4. XRUC4 - Industrial Metaverse and XR devices

Industrial cooperative tasks in a three-dimensional extended reality setting can consist of different virtuality types, which correspond to the immersive degree of the virtual environment.

Augmented Reality (AR) technologies overlays virtual content on the real-world environment, whereas Virtual Reality (VR) fully immerses an end-user into the virtual world. In the industrial setting, oftentimes stakeholders may wish to collaborate during fast-prototyping (or other XR tasks such as factory planning, training, etc) using AR and VR devices, depending on the needs and preference of the end-user. End-users may even wish to not interact with the virtual content in the XR space, but rather just spectate on a simple screen (e.g., Apple iPad tablet with a client application installed on the native iOS operating system). This use case scenario aims to use the HORSE platform to assess the network infrastructure capability which would enable a multi-player session with users who are connected using different XR device mediums (AR, VR, spectator viewing on an Apple iPad iOS tablet). In order to enable this multi-platform approach, the following has to be supported:

- AR → AR + iOS
- AR → VR + iOS
- VR → VR + iOS

The mentioned iOS addition should only be used to view a session, which would not allow interaction with the virtual models, as is possible for the other hardware (e.g., AR HoloLens 2 glasses used in the above scenarios). Additionally, this use case will leverage the features of the HORSE platform, which will enable the above use case scenarios to be deployed while simultaneously maintaining network infrastructure service requirements such as high-bandwidth and low-latency and ensuring end-to-end security measures within the framework of 6G networks.

# 5. HORSE Functional & Non-Functional Requirements

The following table includes a preliminary list of functional requirements of the use cases described previously. The "REQ ID" and the "Name" columns are used to identify the requirement, while the "Description" column provides a more detailed explanation of the requirement. The keywords used in the "Priority" column are to be interpreted as described in RFC 2119 [79]. The "HORSE Module" column links the requirement with the module in the HORSE architecture responsible for implementing the requirement

| REQ ID | Name | Description | Priority | HORSE Module | Use Case |
|---|---|---|---|---|---|
| REQ-F-01 | Multi-device connectivity monitoring | The HORSE platform must monitor the connectivity of devices connected to the managed network | MUST | Smart Monitoring | LTUC1, LTUC2, LTUC3, LTUC4, XRUC1, XRUC2, XRUC4 |
| REQ-F-02 | Multi-device protection | The HORSE platform must monitor the cybersecurity, including authentication, authorization, threat detection and secure connectivity, of the devices on extreme edge connected to the monitored network | MUST | Smart Monitoring | LTUC1, LTUC2, LTUC3, LTUC4, XRUC1, XRUC2, XRUC4 |
| REQ-F-03 | Auditing of messages | The HORSE platform could offer auditing capabilities per subsystem | COULD | IBI, PIL, STO | LTUC1, LTUC2, LTUC3, LTUC4, XRUC1, XRUC2, XRUC4 |
| REQ-F-04 | Auditing of device connections | The HORSE platform must produce auditing logs of devices connecting to the network managed by the platform | MUST | Smart Monitoring | LTUC1 |

| | | | | | |
|---|---|---|---|---|---|
| REQ-F-05 | Unauthorised device attempt detection | The HORSE platform must be able to detect when an unauthorized device attempts to connect to network and stop it from harming the HORSE platform | MUST | Smart Monitoring | LTUC1, LTUC2, LTUC3, LTUC4, XRUC1, XRUC2, XRUC4 |
| REQ-F-06 | Detection of connection error on devices | The HORSE platform must detect connection errors of authorized devices trying to join the HORSE network / slice | MUST | Smart Monitoring | LTUC1, LTUC2, LTUC3, LTUC4, XRUC1, XRUC2, XRUC4 |
| REQ-F-07 | Network Performance monitoring | The HORSE platform must ensure efficient monitoring mechanisms to timely identify network performance degradation regarding reliability, latency, and bandwidth | MUST | Smart Monitoring | LTUC1, LTUC2, LTUC3, LTUC4, XRUC1, XRUC2, XRUC4 |
| REQ-F-08 | Authentication support | The HORSE platform should be able to supervise different authentication and authorization mechanisms (Ie: OAuth2.0, DIDs, digital signatures) | SHOULD | ePEM | LTUC1, LTUC2, LTUC3, LTUC4, XRUC1, XRUC2, XRUC4 |
| REQ-F-09 | Thread detection | The HORSE platform must be able to detect potential threats from external entities with respect to the above authentication and authorization mechanisms | MUST | ePEM | LTUC1, LTUC2, LTUC3, LTUC4, XRUC1, XRUC2, XRUC4 |
| REQ-F-10 | Data integrity | The HORSE platform must be able to verify the integrity of information exchanged between different HORSE modules | SHOULD | RTR, ePEM | LTUC1, LTUC2, LTUC3, LTUC4, XRUC1, XRUC2, XRUC4 |
| REQ-F-11 | Notification | The HORSE platform should be able to notify the network operator about detected threats or foreseen threats | MUST | ePEM, RTR | LTUC1, LTUC2, LTUC3, LTUC4, XRUC1, XRUC2, XRUC4 |
| REQ-F-12 | Detecting of attacks | The HORSE platform should detect attacks (such as DDoS attacks) on the network | SHOULD | ePEM, RTR | LTUC1, LTUC2, LTUC3, LTUC4, XRUC1, XRUC2, XRUC4 |
| REQ-F-13 | Mitigation of attacks | The HORSE platform should propose network and system reconfigurations to mitigate attacks | SHOULD | DTE, PEM | LTUC1, LTUC2, LTUC3, LTUC4, XRUC1, XRUC2, XRUC4 |
| REQ-F-14 | Prediction of attacks | The HORSE platform must implement a methodology to predict an attack | MUST | PEM, EM, DTE | LTUC1, LTUC2, LTUC3, LTUC4, XRUC1, XRUC2, XRUC4 |
| REQ-F-15 | Policy enforcement | The HORSE platforms should be able to enforce reconfiguration of the infrastructure in case of threat or attack detection | SHOULD | RTR, ePEM, | LTUC1, LTUC2, LTUC3, LTUC4, XRUC1, XRUC2, XRUC4 |
| REQ-F-16 | Sandboxing of reconfiguration | The HORSE platform should be able to test and evaluate new configurations before deploying them to the infrastructure | SHOULD | SAN, EM | LTUC1, LTUC2, LTUC3, LTUC4, XRUC1, XRUC2, XRUC4 |
| REQ-F-17 | Persistence of information | The HORSE platform must save the intents entered by the user thought the Intent GUI in a persistent manner | MUST | IBI (Intent Manager) | LTUC1, LTUC2, LTUC3, LTUC4, XRUC1, XRUC2, XRUC4 |

Co-funded by the European Union

| | | | | | |
|---|---|---|---|---|---|
| REQ-F-18 | Reporting of Intent-based decisions | The HORSE platform should report to the network administrator the decisions taken by the intent-based module | SHOULD | IBI (Learning and Reasoning / Dashboard) | LTUC1, LTUC2, LTUC3, LTUC4, XRUC1, XRUC2, XRUC4 |
| REQ-F-19 | Anomaly detection | The HORSE platform must provide the appropriate mechanisms for anomaly detection in the transmitted messages | MUST | PEM | LTUC1, LTUC2, LTUC3, LTUC4, XRUC1, XRUC2, XRUC4 |
| REQ-F-20 | 6G components modelling | The HORSE platform should identify and model 6G components. | SHOULD | EM | LTUC1, LTUC2, LTUC3, LTUC4, XRUC1, XRUC2, XRUC4 |
| REQ-F-21 | Attack modelling | The HORSE platform must be able to model attacks, and its impact. | MUST | EM, PEM | LTUC1, LTUC2, LTUC3, LTUC4, XRUC1, XRUC2, XRUC4 |
| REQ-F-22 | Awareness of slices | The HORSE platform must be aware of slices in the networks (level of isolation, shared elements) | MAY | Slice Manager | LTUC1, LTUC2, LTUC3, LTUC4, XRUC1, XRUC2, XRUC4 |
| REQ-F-23 | Access management | The user must be able to define and then the HORSE platform must enforce access policies in real time and ensure that the information is only available to authorized users | MUST | PAG, IBI, | LTUC1, LTUC2, LTUC3, LTUC4, XRUC1, XRUC2, XRUC4 |
| REQ-F-24 | Secure interaction with the exposure functions of the network | The HORSE platform should be able to securely exchange data and commands from/to the network | SHOULD | RAN, CORE | LTUC1, LTUC2, LTUC3, LTUC4, XRUC1, XRUC2, XRUC4 |
| REQ-F-25 | Granularity of the access | The HORSE platform should be able to support different roles in accessing the system | SHOULD | IBI (Dashboard), PAG | LTUC1, LTUC2, LTUC3, LTUC4, XRUC1, XRUC2, XRUC4 |
| REQ-F-26 | Multi-tier orchestration | The HORSE platform should be able to manage services in a multi-tier environment, comprising cloud, edge and far edge | SHOULD | SM, RTR, ePEM | LTUC1, LTUC2, LTUC3, LTUC4, XRUC1, XRUC2, XRUC4 |
| REQ-F-27 | Notification of provisioning | The HORSE platform must notify the user about reconfiguration of the network to mitigate or avoid an attack | MUST | IBI, EM, PEM | LTUC1, LTUC2, LTUC3, LTUC4, XRUC1, XRUC2, XRUC4 |
| REQ-F-28 | Real-time monitoring of attacks | The HORSE platform must continuously monitor the network for possible attacks and notify the user about them | MUST | PEM | LTUC1, LTUC2, LTUC3, LTUC4, XRUC1, XRUC2, XRUC4 |
| REQ-F-29 | Policies definition | The HORSE platform should provide a way to define policies and action to be performed in the network when some certain conditions are met | SHOULD | IBI, RTR, PEM | LTUC1, LTUC2, LTUC3, LTUC4, XRUC1, XRUC2, XRUC4 |
| REQ-F-30 | Configuration of security level per slice | The HORSE platform must support configuration of security level in a per-slice granularity | MUST | ePEM | LTUC1, LTUC2, LTUC3, LTUC4, XRUC1, XRUC2, XRUC4 |
| REQ-F-31 | Use of anonymized data for IA training | The HORSE platform should use anonymized data to train AI model to detect threats and attacks | SHOULD | DTE | LTUC1, LTUC2, LTUC3, LTUC4, XRUC1, XRUC2, XRUC4,. |

| | | | | | |
|---|---|---|---|---|---|
| REQ-F-32 | AI-based policies definition | The HORSE platform should be able to define set of optimum policies using AI/ML models to guarantee the system security against potential attacks | SHOULD | DTE | LTUC1, LTUC2, LTUC3, LTUC4, XRUC1, XRUC2, XRUC4,. |
| REQ-F-33 | Reproducibility and repeatability of certain behaviours of digital twin in the network | The HORSE Digital Twin should be able to consistently repeat specific experiments, and to incorporate controlled variations to experiments execution as requested by its users. | SHOULD | SAN | LTUC1, LTUC2, LTUC3. LTUC4. XRUC1, XRUC2, XRUC3, XRUC4 |
| REQ-F-34 | Different granularity of control plane functions for Digital twin | The HORSE Digital Twin should be capable of deploying different network functions to test them independently or to model whole functionality sets or planes as a single entity, according to specific experiment. | SHOULD | SAN | LTUC1, LTUC2, LTUC3. LTUC4. XRUC1, XRUC2, XRUC3, XRUC4 |
| REQ-F-35 | Data anonymization | The HORSE platform must execute data anonymization operations on collected data assets | MUST | PAG | LTUC1, LTUC2, LTUC3, LTUC4, XRUC1, XRUC2, XRUC4 |
| REQ-F-36 | Data encryption | The HORSE platform must support end-to-end data encryption for data in transit | MUST | PAG | LTUC1, LTUC2, LTUC3, LTUC4, XRUC1, XRUC2, XRUC4 |
| REQ-F-37 | Observability | The HORSE platform should allow the user to monitor the status (successful or failed execution) and view an incident summary of all AI pipelines | SHOULD | PAG, DTE | LTUC1, LTUC2, LTUC3, LTUC4, XRUC1, XRUC2, XRUC4 |
| REQ-F-38 | Data retention | The user should be able to define and then the HORSE platform should execute data retention operations (e.g., automated deletion after a certain due date) on collected data assets | SHOULD | PAG | LTUC1, LTUC2, LTUC3, LTUC4, XRUC1, XRUC2, XRUC4 |
| REQ-F-39 | Data ingestion 1 | The HORSE platform must allow the ingestion of data at rest (for example, from a file or from an API) | | Smart Monitoring | LTUC1, LTUC2, LTUC3, LTUC4, XRUC1, XRUC2, XRUC4 |
| REQ-F-40 | Data ingestion 2 | The HORSE platform must allow the ingestion of real-time data (for example, streaming data) | | Smart Monitoring | LTUC1, LTUC2, LTUC3, LTUC4, XRUC1, XRUC2, XRUC4 |
| REQ-F-41 | Data pre-processing | The user must be able to define data pre-processing rules (for handling outliers, for handling missing data values, etc.) on the data assets and the HORSE platform must pre-process the collected data assets according to these rules | | Smart Monitoring | LTUC1, LTUC2, LTUC3, LTUC4, XRUC1, XRUC2, XRUC4 |
| REQ-F-42 | Network and service status information | The PIL must be able to periodically gather information about the status of the network and running services | MUST | PIL | LTUC1, LTUC2, LTUC3, LTUC4, XRUC1, XRUC2, XRUC4 |
| REQ-F-43 | Threat Detection Time | The Horse platform must be able to provide early attack detection within three ROPs | MUST | Threat Detector and Mitigation Engine (PEM) | LTUC1, LTUC2, LTUC3. LTUC4. |

| | | | | | |
|---|---|---|---|---|---|
| | | (Note: The ROP period with which monitoring data are collected from the network is dimensioned according the network size. A typical value is in the order of some minutes) | | | XRUC1, XRUC2, XRUC3, XRUC4 |
| REQ-F-44 | Threat Detection Rate | The detection rate, i.e., the number of successfully detected attacks over the total, should be above 90%, considering the application of ML and the SoA benchmarks | SHOULD | Threat Detector and Mitigation Engine (PEM) | LTUC1, LTUC2, LTUC3. LTUC4. XRUC1, XRUC2, XRUC3, XRUC4 |

*Table 1: List of functional requirements*

The next table includes the list of non-functional requirements. Each non-functional requirement is identified by its requirement id (REQ ID) and name. The column "Description" is used to detail the non-functional requirement, while the "Priority" column is to be interpreted as described in RFC 2119 [79].

| REQ ID | Name | Description | Priority | Use Case |
|---|---|---|---|---|
| REQ-NF-01 | Geographic dispersion support | The HORSE platform must consider the geographic dispersion of elements and devices connecting to the network. | MUST | LTUC1, LTUC2, LTUC3, LTUC4, XRUC2 |
| REQ-NF-02 | Scalability | The HORSE platform should support expansion and scalability of the systems. | SHOULD | |
| REQ-NF-03 | Centralized data collection | The HORSE platform should collect data about the network performance using a centralized endpoint | SHOULD | |
| REQ-NF-04 | User friendliness | The HORSE platform should offer intuitive user-friendly interfaces | SHOULD | |
| REQ-NF-05 | Consistent interfaces | The HORSE platform should offer consistent interfaces among its different modules | SHOULD | |
| REQ-NF-06 | Decentralized management | The HORSE platform should support decentralized management of system modules (not all modules hosted in the same datacentre) | SHOULD | |
| REQ-NF-07 | Installability of the platform | The HORSE platform should be installable in servers running at the cloud or at servers running at the edge or the network (e.g., on premises servers) | SHOULD | |
| REQ-NF-08 | Platform support | The HORSE platform components must support installation on the Linux platform | MUST | |
| REQ-NF-09 | Web access | The components of the HORSE platform should be accessible via web technologies | SHOULD | |
| REQ-NF-10 | Support of lightweight virtualization | The HORSE platform should be based on containerized software blocks in order to support service mobility in a timely manner | SHOULD | |
| REQ-NF-11 | Compliance to legal legislation | The HORSE platform must comply with the legal framework with respect to information dissemination | MUST | |
| REQ-NF-12 | Service mobility | The HORSE Platform must support device and service mobility. | SHOULD | |
| REQ-NF-13 | | The IA methods employed by the HORSE platform must adhere to ethical principles and values | MUST | |
| REQ-NF-14 | Context awareness | The HORSE platform must have access to status information about network traffic, network status and services | MUST | ALL |

*Table 2: List of non-functional requirements*

# 6. HORSE Network Services and Threats

This section provides a description of the 6G network services and threats that the HORSE project will consider. The 6G network services requirements will be defined in the network slicing, API exposure, and AI data training research areas, while the HORSE 6G network threats will include DoS, data tampering and network congestion.

## 6.1. 6G Services considered in HORSE

In the HORSE project, 6G network services will cover at least the following three areas: (i) network slicing isolation, (ii) API exposure, and (iii) AI data training.

### 6.1.1. Network slicing isolation

Network slicing is a key feature for 5G networks introduced by 3GPP from Release 15 [80], posing new security challenges. These challenges have been tentatively addressed following the security-by-design principle for the management of common infrastructure and resources sharing amongst network slices. In further releases [81], new security requirements have been identified related to privacy and access control, including: i) the secure management for network slices, and; ii) network slice specific authentication and authorization (NSSAA). In this scenario it is absolutely clear that network slicing security will continue to evolve incorporating new security features, as for example those related to the management of devices connected to various slices at the same time for diversified service access to avoid leakage. This has made network slices isolation to become a key issue to enhance security in 5G networks. In 6G, it is expected that network slicing will be used at all levels, i.e., physical, network and service. Thus, end-to-end isolation will enhance 6G security significantly, but it is still an open issue that will require the proper identification of a set of well-defined specifications.

### 6.1.2. API exposure

APIs exposure at the edge is ever gaining importance in 5G networks motivated by the need to enrich existing services with improved security and performance, as well as to enable the development of new ones. Indeed, highly demanding 5G services (e.g., those requiring low latency or supported by massive IoT deployments) may benefit from 5G network capabilities, including for example high performance, an efficient data management (collection, processing, ingestion), or a proper device location, to offer high levels of quality of service, even for extremely demanding services.

The 5GPP architecture working group in the "The 6G Architecture Landscape European perspective" white paper [82], pointed out API exposure as a key research area, being a fundamental component in the global architecture to support and facilitate applications to interact with the network.

### 6.1.3. AI data training

The ever-growing need for data to be used to generate and train AI models is nowadays a key driver in research to which 5G/6G is not blind to. Many of the well described challenges 6G must face in the coming years may be addressed by developing AI-assisted solutions. When dealing with security provisioning AI data training will undoubtedly play a key role in defining predictive models that may notably contribute to design proactive and more automated security strategies. However, although many aspects related to how data is collected, ingested, stored and preserved yet remain as challenges, it is with no doubt that the development of customized AI training models for 5G/6G network scenarios would contribute to the design of novel services, where security will become a key pillar fostering the creation of novel services or even business models.

## 6.2. 6G threats considered in HORSE

HORSE identifies possible threats for the two use cases included in the project in order to set a clear specification of requirements, a complete set of functionalities as well as the required needs for preliminary testing and validating the outcome of the project.

### 6.2.1. Secure Smart LRT Systems Use Case

The threats HORSE considers to be addressed in the Secure Smart LRT Systems Use Case, are Denial of Service (DoS) and data tampering. The rationale behind this assessment is motivated by the previous work done by ENISA in the last Threat Landscape Report [78] for the transport sector, identifying these two treats as two of the main threats targeting the railway sector. Indeed, most of the last year attacks in the railway sector targeted their IT systems causing disruptions in passenger services, display boards, surveillance systems, etc.

### 6.2.2. Remote Rendering to Power XR Industrial Use Case (R$^2$2XRI)

The Remote Rendering to Power XR Industrial Use Case (R$^2$2XRI) aims to leverage the resilience and security functionalities to be provided by the HORSE project. These functionalities will ensure an unhindered data exchange flow during XR experiences which allows maintained operational efficiency in the industrial setting. To assess the efficacy of the envisioned security functionalities, a dedicated network threat will be simulated during the execution of this use case. Usage of the XR technologies here consists of a XR application hosted on a server which sends application streams to XR devices (e.g., HoloLens 2 AR glasses) and in return receives back sensor data. The reciprocal data exchange between XR device and XR application requires quality connectivity to ensure timely packet flow and permissive levels of transmission to avoid congestion. The latter of can be highly detrimental for end-users, as visualization and interaction with XR content is significantly impacted if network connectivity is highly congested and can even cease altogether. The simulated network threat in this use case will saturate the XR network connectivity to drive a low bit rate for data exchange. This threat scenario will be used to assess the extent to which HORSE can respond and effectively resolve such an attack within the context of XR experiences.

# 7. HORSE AI Data Management

## 7.1. Data Ingestion Mechanisms

The ingestion of various types of data, such as 5G/6G network traffic, lightweight rail train operation data, command centre operation data, remote application rendering data, is important as it constitutes the initial step towards their insightful analysis in HORSE. In this context, Data Ingestion is responsible for the ingestion of datasets prior to their storage in the HORSE platform and implements the following functionalities:

- **Step-by-step definition of the data ingestion process:** Data Ingestion provides a user-friendly interface that enables the data providers to specify the desired settings for the ingestion of their datasets. A large group of configuration specifications is offered, including for example, harvesting options, authentication details, and retrieval schedule.
- **Flexible management of data ingestion:** Data Ingestion enables the creation of a configuration file where all data ingestion settings for a dataset are saved, as defined by its data provider; these settings may also be changed without restriction until the finalization of the configuration file. Once the configuration file is finalized, updates are permitted only for certain settings so that the data ingestion process runs smoothly and consistently.
- **Retrieval of datasets as files:** Data Ingestion enables the retrieval and uploading of different types of files, *for example* (a) tabular (e.g. CSV), (b) non-tabular (e.g. XML, JSON) and (c) others (e.g. images). The actual file types will be identified by the project's Use Cases. It also supports the uploading of samples of the datasets to facilitate the creation of the configuration files.
- **Retrieval of datasets through APIs:** Data Ingestion supports the ingestion of data through external 3rd party APIs by offering a detailed interface for the definition of API specifications, such as the method, the path, the query parameters, and the retrieval settings of the API connection. The actual need for ingestion through APIs and the APIs definition will be identified by the project's Use Cases.
- **Selection of certain data to be stored by the data provider:** Data Ingestion enables the data providers to select which data they want to store in HORSE. As a result, only the selected data of a dataset will undergo further check-in processing, while the rest of the data will be discarded and not stored in HORSE.
- **Data update management:** Data Ingestion facilitates the update of datasets through uploading of additional files that append the already stored datasets in HORSE. These files should have the same data structure as the initial uploaded files and the update is performed according to the already defined pre-processing rules of the configuration file used for the data collection of the initial file.
- **Generation of data ingestion status messages:** Data Ingestion shall communicate the status of its data ingestion jobs through feedback messages (e.g., if no API results were retrieved due to server or user errors along with their respective code).

## 7.2. Data Management Procedures

The provision of harmonized, interoperable, and consistent datasets is critical to enable data providers to better understand their data and perform analytics tasks by combining datasets from different sources in an easier manner. Data Management procedures provide:

- **Mapping of the ingested data to a respective HORSE data model:** Data Management aligns the ingested datasets to a common HORSE data model to enhance their interoperability. Part of the mapping process is performed in a semiautomated way; as a second step, the data provider is allowed to review the automatic mappings through a user-friendly interface and decide whether these suggestions will be maintained, modified or deleted.

- **Data pre-processing:** Data Management ensures the data quality of the available assets, which is of crucial role for ML models since it impacts the accuracy and generalizability of the model. Errors in data can be due to various factors. They can be the results of imperfections in the data acquisition system or deliberate attacks aiming to poison the data. Both types of errors can harm the performance of the current model and specific measures should be taken in each case due to the structural differences between them. For minimising imperfections in the data acquisition system, possible techniques include:

  - Data profiling and statistical checks: Analyse statistical properties and data distribution to locate distribution drifts related to erroneous samples.
  - Outlier detection: Utilize outlier detection algorithms to identify anomalies errors in the data.
  - Domain specific rules (including human expert knowledge): Ingest prior knowledge in the form of filters to exclude datapoints not complying to this knowledge.

  Another possible type of harmful data is input generated by adversarial attacks. Data management should exclude such data, before ML training, in order to create robust and secure AI models. In adversarial attacks, the attacker forces the model to produce incorrect outputs by slightly perturbating the input in such a way (often unnoticeable by humans) that elicits misclassification of the model. Another form of adversarial attack is data poisoning, where the attacker manipulates the input data used in ML training in order to decrease the performance of the model, shift the classification outcome of a specific sample to the wrong class, increase training time, etc. [83]

  Data management uses the following mitigation methods to tackle these threats [84]:

  - Adversarial training: synthetic/adversarial samples are created from the original dataset to be used also in training. Therefore, using augmented data, the model’s resilience is increased.
  - Input validation and filtering: validation checks are performed on the input to identify unexpected patterns and abnormalities. Flagged adversarial examples are then discarded based on specific criteria and thresholds.
  - Pre-processing techniques: a simple process the involves normalizing the inputs prior to feeding the ML model like normalization, dimensionality reduction or feature scaling.

- **Generation of data handling status messages:** Data Management is designed with observability in mind for the constant and efficient monitoring of the data pipelines’ execution to timely detect potential problems by identifying any deviations from their corresponding configuration and efficiently communicating the status of its data management processes and any errors encountered through a number of feedback messages (e.g., the number of “rows” in the data that were dropped due to inconsistent data types)
- **Data retention and disposal:** Data Management allows the data provider to assess whether data is needed and if it can be deleted, by examining the different policies surrounding the usage and the access of the data, in conjunction with the required by the legislation actions that should be performed (as for example GDPR clauses, or legislation regarding data retention for auditing purposes), and suggests the most optimal methods for the erasure of these artefacts.

- **Storage of data:** Once the data are appropriately ingested and handled, the processed data along with the processed sample are forwarded to the storage component.
- **Storage of assets' metadata:** Along with the processed data assets, the accompanying metadata are also forwarded to the storage component, creating the appropriate links between the stored data and their metadata information.

## 7.3. Description of Use Case Datasets

In order to ensure compliance with GDPR regulations and protect personal data, the datasets used in this project will undergo anonymization processes. We prioritize the privacy and security of individuals involved, and no personal data will be utilized. The detailed data management plan will be included in deliverable D1.2, outlining the strategies and procedures for handling and storing the datasets securely. It is important to note that the description of the datasets will be formulated in the upcoming deliverable D2.3. This approach is taken due to ongoing architecture discussions that are yet to be finalized, ensuring that the dataset description aligns accurately with the project's objectives.

## 7.4. 5G/6G Network Traffic Data

The HORSE project will consider 5G network traffic public datasets, as well as realistic cyber defense datasets.

Cyber defense datasets enhance Intrusion Detection Systems (IDSs) due to its potential in detecting zero-day attacks. They consist of network traffic, including benign data flows and cyberattacks. They will speed-up the adoption of ML-based IDSs in real-world applications, as they will facilitate the training, tuning and evaluation prior to deployment. The most recent public datasets are the following:

- The **5G-NIDD dataset** [85] is a labeled NID dataset built on a real 5G test network. It consists of a combination of benign 5G network flows and attack traffic. The attacks considered are DoS, DDoS and port scan attacks.
- The **IoT-23 dataset** [86] is a labeled dataset with malicious and benign IoT network traffic. It has been developed by the Stratosphere Laboratory of the CTU University. This large dataset consists of real data and labeled IoT malware infections, consisting 3 captures for benign IoT devices and 20 malware captures.
- The **CSE-CIC-IDS2018 dataset** [87] has been developed by the Communications Security Establishment (CSE) and the Canadian Institute for Cybersecurity (CIC). This initiative enables the generation of diverse datasets based on user profiles. Each of the generated datasets contains a unique set of features according to the evaluation domain. The last version of the dataset includes seven different attacks: Bruteforce, DoS, DDoS, infiltration, botnet, and Web attack infiltration of the network from inside. The dataset includes network traffic flows and systems logs, along with 80 features extracted using CICFlowMeter.
- The **BoT-IoT dataset** [88] has been created by the Australian Center for Cyber Security (ACCS), developing a realistic network environment in the Cyber Range Lab of UNSW Canberra. It comprises over 73 millions of records of network activity in a simulated IoT environment, including both normal and several cyberattack traffic flows. The dataset has 46 features and 5 different output classes, one for normal traffic and four for the different types of attacks (DoS, DDoS, reconnaissance, and information theft).
- The **UNSW-NB15 dataset** [89] is also developed by the ACCS in collaboration with researchers worldwide. It consists of normal and synthesized attack activities in a simulated IoT environment. This dataset contains normal traffic and nine different types

of attacks (generic, exploits, fuzzers, DoS, reconnaissance, analysis, backdoor, shellcode, and worms). The UNSW-NB15 dataset has 49 different features and it comprises over 2 million records stored in four different CSV files.

# 8. Conclusions

In this deliverable, we have presented an overview of the HORSE landscape, which focuses on the development of Holistic, Omnipresent, Resilient Services for future 6G Wireless and Computing Ecosystems. Our preliminary work has explored various research pillars and conducted a state-of-the-art analysis to identify key areas of interest and technological advancements.

Within the realm of security, we have examined the implications of the 6G world and discussed risks and threats that need to be addressed. Furthermore, we have emphasized the importance of threat identification, characterization, and modelling to enhance security measures in the 6G ecosystem.

Networking capabilities are a crucial aspect of the HORSE project, and we have delved into several key areas within this domain. Our analysis has explored network exposure capabilities beyond 5G, focusing on energy efficiency and the design of digital twins. Additionally, we have highlighted the significance of the physical layer in 6G networks to ensure robust and reliable connectivity.

Artificial Intelligence (AI) plays a pivotal role in shaping the future of wireless and computing ecosystems. Therefore, we have investigated AI-enabled solutions for enhancing security in 6G and mitigating threats. Moreover, intent-based networking has emerged as a promising approach to optimize network operations and improve overall efficiency. In the same manner we analysed their capabilities in various points of the text.

To demonstrate the practical implications, the HORSE project, is implementing two use cases which have been also presented in this deliverable. The first use case, Secure Smart LRT Systems (SS-LRT), showcases the application of HORSE infrastructure in ensuring the security and smooth operation of smart public transportation systems. The second use case focuses on extended reality (XR) and highlights the benefits of rendering XR content in local networks, enabling fast prototyping sessions, multi-user experiences, and industrial applications.

Building upon these use cases and research pillars, we have outlined the functional and non-functional requirements that will guide the technical work of the HORSE project. These requirements encompass various aspects such as security, networking, AI, and data management to ensure the development of holistic and resilient services.

Furthermore, considering the importance of data in driving advancements, we have discussed the data ingestion mechanisms and management procedures within the HORSE project. Additionally, we have described the potential datasets associated with the use cases, providing valuable insights into the types of data and workflows involved.

Moving forward, this preliminary work sets the foundation for the technical implementation of the HORSE project. The insights gained from the state-of-the-art analysis, use cases, and requirements will serve as valuable references for the development of cutting-edge technologies and policies that align with the goals of the project.

As we embark on the next phase of this research endeavour, we are confident that the HORSE project will contribute significantly to the advancement of 6G wireless and computing ecosystems.

# References

[1] “European Vision for the 6G Network Ecosystem ‹ 5G-PPP.” https://5g-ppp.eu/european-vision-for-the-6g-network-ecosystem/ (accessed Jun. 25, 2023).

[2] W. Jiang, B. Han, M. A. Habibi, and H. D. Schotten, “The Road Towards 6G: A Comprehensive Survey,” *IEEE Open J. Commun. Soc.*, vol. 2, pp. 334–366, 2021, doi: 10.1109/OJCOMS.2021.3057679.

[3] F. Mademann, “The 5G system architecture,” *J. ICT Stand.*, pp. 77–86, 2018.

[4] B. Han, V. Gopalakrishnan, L. Ji, and S. Lee, “Network function virtualization: Challenges and opportunities for innovations,” *IEEE Commun. Mag.*, vol. 53, no. 2, pp. 90–97, 2015.

[5] D. Jones, C. Snider, A. Nassehi, J. Yon, and B. Hicks, “Characterising the Digital Twin: A systematic literature review,” *CIRP J. Manuf. Sci. Technol.*, vol. 29, pp. 36–52, 2020.

[6] W. Saad, M. Bennis, and M. Chen, “A vision of 6G wireless systems: Applications, trends, technologies, and open research problems,” *IEEE Netw.*, vol. 34, no. 3, pp. 134–142, 2019.

[7] “A game plan for quantum computing | McKinsey.” https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/a-game-plan-for-quantum-computing (accessed Jun. 25, 2023).

[8] R. J. McEliece, “A public-key cryptosystem based on algebraic,” *Coding Thv*, vol. 4244, pp. 114–116, 1978.

[9] J. Hoffstein, J. Pipher, and J. H. Silverman, “NTRU: A ring-based public key cryptosystem,” in *International algorithmic number theory symposium*, Springer, 1998, pp. 267–288.

[10] I. T. L. Computer Security Division, “Post-Quantum Cryptography | CSRC | CSRC,” *CSRC | NIST*, Jan. 03, 2017. https://csrc.nist.gov/projects/post-quantum-cryptography (accessed Jun. 25, 2023).

[11] “Specification # 33.841.” https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3422 (accessed Jun. 25, 2023).

[12] S. A. Abdel Hakeem, H. H. Hussein, and H. Kim, “Security requirements and challenges of 6G technologies and applications,” *Sensors*, vol. 22, no. 5, p. 1969, 2022.

[13] P. Porambage, G. Gür, D. P. M. Osorio, M. Livanage, and M. Ylianttila, “6G security challenges and potential solutions,” in *2021 Joint European Conference on Networks and Communications & 6G Summit (EuCNC/6G Summit)*, IEEE, 2021, pp. 622–627.

[14] M. Wang, T. Zhu, T. Zhang, J. Zhang, S. Yu, and W. Zhou, “Security and privacy in 6G networks: New areas and new challenges,” *Digit. Commun. Netw.*, vol. 6, no. 3, pp. 281–291, 2020.

[15] M. Liyanage, A. Braeken, S. Shahabuddin, and P. Ranaweera, “Open RAN security: Challenges and opportunities,” *J. Netw. Comput. Appl.*, vol. 214, p. 103621, 2023.

[16] “Cybersecurity of Open Radio Access Networks | Shaping Europe’s digital future,” May 11, 2022. https://digital-strategy.ec.europa.eu/en/library/cybersecurity-open-radio-access-networks (accessed Jun. 25, 2023).

[17] P. V. Bro, “A system for detecting network intruders in real-time,” in *Proc. 7th USENIX security symposium*, 1998.

[18] M. Roesch, “Lightweight intrusion detection for networks,” in *Proceedings of LISA*, 2005.

[19] C. Chio and D. Freeman, *Machine learning and security: Protecting systems with data and algorithms*. O'Reilly Media, Inc., 2018.

[20] M. Dekker and I. Bakatsis, "Cybersecurity Incident Taxonomy," 2018.

[21] N. Long and R. Thomas, "Trends in denial of service attack technology," *CERT Coord. Cent.*, vol. 648, p. 651, 2001.

[22] D. Moore, "The spread of the sapphire/slammer worm," *Httpwww Cs Berkeley Edu~ Nweaversapphire*, 2003.

[23] "Machine Learning in Cybersecurity Course - Part 1: Core Concepts and Examples," *NopSec*, Mar. 25, 2019. https://www.nopsec.com/blog/part-1-ml-in-cybersecurity/ (accessed Jun. 25, 2023).

[24] V. Saini, Q. Duan, and V. Paruchuri, "Threat modeling using attack trees," *J. Comput. Sci. Coll.*, vol. 23, no. 4, pp. 124–131, 2008.

[25] Y. Chen, B. Boehm, and L. Sheppard, "Value driven security threat modeling based on attack path analysis," in *2007 40th Annual Hawaii International Conference on System Sciences (HICSS'07)*, IEEE, 2007, pp. 280a–280a.

[26] "Network programmability in 5G: an invisible goldmine for service providers and industry." https://www.ericsson.com/en/blog/2019/1/network-programmability---in-5g-an-invisible-goldmine-for-service-providers-and-industry (accessed Jun. 25, 2023).

[27] "'NetApp: Opening up 5G and beyond networks' White Paper," *EVOLVED-5G*, Dec. 14, 2022. https://evolved-5g.eu/2022/12/14/netapp-opening-up-5g-and-beyond-networks-white-paper/ (accessed Jun. 25, 2023).

[28] "NWDAF: Automating the 5G network with machine learning and data analytics." https://inform.tmforum.org (accessed Jun. 25, 2023).

[29] "ETSI TS 133 521 V17.1.0 (2022-05) - 5G; 5G Security Assurance Specification (SCAS);Network Data Analytics Function (NWDAF) (3GPP TS 33.521 version 17.1.0 Release 17)," *iTeh Standards*. https://standards.iteh.ai/catalog/standards/etsi/2c528851-5a69-4889-a4e5-5dc31b138bac/etsi-ts-133-521-v17-1-0-2022-05 (accessed Jun. 25, 2023).

[30] K. J. Rahman *et al.*, "Challenges, applications and design aspects of federated learning: A survey," *IEEE Access*, vol. 9, pp. 124682–124700, 2021.

[31] J. C. Cisneros, S. Yangui, S. E. P. Hernández, and K. Drira, "A survey on distributed NFV multi-domain orchestration from an algorithmic functional perspective," *IEEE Commun. Mag.*, vol. 60, no. 8, pp. 60–65, 2022.

[32] F. Z. Yousaf, M. Bredel, S. Schaller, and F. Schneider, "NFV and SDN—Key technology enablers for 5G networks," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 11, pp. 2468–2478, 2017.

[33] H. Ibn–khedher, M. Hadji, and A. E. Kamal, "Placement, Routing and Scheduling Optimizations in Cloud-RAN," in *2021 IEEE Global Communications Conference (GLOBECOM)*, IEEE, 2021, pp. 01–06.

[34] R. Zhou *et al.*, "Online task offloading for 5G small cell networks," *IEEE Trans. Mob. Comput.*, vol. 21, no. 6, pp. 2103–2115, 2020.

[35] M. B. Janjua and H. Arslan, "A Survey of Symbiotic Radio: Methodologies, Applications, and Future Directions," *Sensors*, vol. 23, no. 5, p. 2511, 2023.

[36] H. Ahmadi, A. Nag, Z. Khar, K. Sayrafian, and S. Rahardja, "Networked twins and twins of networks: An overview on the relationship between digital twins and 6G," *IEEE Commun. Stand. Mag.*, vol. 5, no. 4, pp. 154–160, 2021.

[37] N. P. Kuruvatti, M. A. Habibi, S. Partani, B. Han, A. Fellan, and H. D. Schotten, “Empowering 6G Communication Systems With Digital Twin Technology: A Comprehensive Survey,” *IEEE Access*, 2022.

[38] S. Vakaruk, A. Mozo, A. Pastor, and D. R. López, “A digital twin network for security training in 5G industrial environments,” in *2021 IEEE 1st International Conference on Digital Twins and Parallel Intelligence (DTPI)*, IEEE, 2021, pp. 395–398.

[39] M. Matthaiou, O. Yurduseven, H. Q. Ngo, D. Morales-Jimenez, S. L. Cotton, and V. F. Fusco, “The road to 6G: Ten physical layer challenges for communications engineers,” *IEEE Commun. Mag.*, vol. 59, no. 1, pp. 64–69, 2021.

[40] M. Giordani, M. Polese, M. Mezzavilla, S. Rangan, and M. Zorzi, “Toward 6G networks: Use cases and technologies,” *IEEE Commun. Mag.*, vol. 58, no. 3, pp. 55–61, 2020.

[41] Y. Liu *et al.*, “Reconfigurable intelligent surfaces: Principles and opportunities,” *IEEE Commun. Surv. Tutor.*, vol. 23, no. 3, pp. 1546–1577, 2021.

[42] E. Basar, M. Di Renzo, J. De Rosny, M. Debbah, M.-S. Alouini, and R. Zhang, “Wireless communications through reconfigurable intelligent surfaces,” *IEEE Access*, vol. 7, pp. 116753–116773, 2019.

[43] E. Basar, “Reconfigurable intelligent surface-based index modulation: A new beyond MIMO paradigm for 6G,” *IEEE Trans. Commun.*, vol. 68, no. 5, pp. 3187–3196, 2020.

[44] E. Basar, I. Yildirim, and F. Kilinc, “Indoor and outdoor physical channel modeling and efficient positioning for reconfigurable intelligent surfaces in mmWave bands,” *IEEE Trans. Commun.*, vol. 69, no. 12, pp. 8600–8611, 2021.

[45] L. Mucchi *et al.*, “Physical-layer security in 6G networks,” *IEEE Open J. Commun. Soc.*, vol. 2, pp. 1901–1914, 2021.

[46] V. Kumar, M. F. Flanagan, D. W. K. Ng, and L.-N. Tran, “On the secrecy rate under statistical QoS provisioning for RIS-assisted MISO wiretap channel,” in *2021 IEEE Global Communications Conference (GLOBECOM)*, IEEE, 2021, pp. 1–6.

[47] Z. Zhang, C. Zhang, C. Jiang, F. Jia, J. Ge, and F. Gong, “Improving physical layer security for reconfigurable intelligent surface aided NOMA 6G networks,” *IEEE Trans. Veh. Technol.*, vol. 70, no. 5, pp. 4451–4463, 2021.

[48] N. Mensi, D. B. Rawat, and E. Balti, “Physical layer security for V2I communications: Reflecting surfaces vs. relaying,” in *2021 IEEE Global Communications Conference (GLOBECOM)*, IEEE, 2021, pp. 01–06.

[49] Y. Siriwardhana, P. Porambage, M. Liyanage, and M. Ylianttila, “AI and 6G security: Opportunities and challenges,” in *2021 Joint European Conference on Networks and Communications & 6G Summit (EuCNC/6G Summit)*, IEEE, 2021, pp. 616–621.

[50] B. Mao, Y. Kawamoto, and N. Kato, “AI-based joint optimization of QoS and security for 6G energy harvesting Internet of Things,” *IEEE Internet Things J.*, vol. 7, no. 8, pp. 7032–7042, 2020.

[51] R. Dong, B. Wang, and K. Cao, “Deep learning driven 3D robust beamforming for secure communication of UAV systems,” *IEEE Wirel. Commun. Lett.*, vol. 10, no. 8, pp. 1643–1647, 2021.

[52] V.-L. Nguyen, P.-C. Lin, B.-C. Cheng, R.-H. Hwang, and Y.-D. Lin, “Security and privacy for 6G: A survey on prospective technologies and challenges,” *IEEE Commun. Surv. Tutor.*, vol. 23, no. 4, pp. 2384–2428, 2021.

[53] S. Shen, C. Yu, K. Zhang, J. Ni, and S. Ci, “Adaptive and dynamic security in AI-empowered 6G: From an energy efficiency perspective,” *IEEE Commun. Stand. Mag.*, vol. 5, no. 3, pp. 80–88, 2021.

[54] D. Je, J. Jung, and S. Choi, “Toward 6G security: technology trends, threats, and solutions,” *IEEE Commun. Stand. Mag.*, vol. 5, no. 3, pp. 64–71, 2021.

[55] E. Rodriguez, B. Otero, N. Gutierrez, and R. Canal, “A survey of deep learning techniques for cybersecurity in mobile networks,” *IEEE Commun. Surv. Tutor.*, vol. 23, no. 3, pp. 1920–1955, 2021.

[56] A. Afaq, N. Haider, M. Z. Baig, K. S. Khan, M. Imran, and I. Razzak, “Machine learning for 5G security: Architecture, recent advances, and challenges,” *Ad Hoc Netw.*, vol. 123, p. 102667, 2021.

[57] O. Jullian, B. Otero, E. Rodriguez, N. Gutierrez, H. Antona, and R. Canal, “Deep-Learning Based Detection for Cyber-Attacks in IoT Networks: A Distributed Attack Detection Framework,” *J. Netw. Syst. Manag.*, vol. 31, no. 2, p. 33, 2023.

[58] H. Sedjelmaci, “Attacks detection approach based on a reinforcement learning process to secure 5g wireless network,” in *2020 IEEE International Conference on Communications Workshops (ICC Workshops)*, IEEE, 2020, pp. 1–6.

[59] S. Jayasinghe, Y. Siriwardhana, P. Porambage, M. Liyanage, and M. Ylianttila, “Federated learning based anomaly detection as an enabler for securing network and service management automation in beyond 5g networks,” in *2022 Joint European Conference on Networks and Communications & 6G Summit (EuCNC/6G Summit)*, IEEE, 2022, pp. 345–350.

[60] S. Kim, K.-J. Park, and C. Lu, “A survey on network security for cyber–physical systems: From threats to resilient design,” *IEEE Commun. Surv. Tutor.*, vol. 24, no. 3, pp. 1534–1573, 2022.

[61] S. Dong, Y. Xia, and T. Peng, “Network abnormal traffic detection model based on semi-supervised deep reinforcement learning,” *IEEE Trans. Netw. Serv. Manag.*, vol. 18, no. 4, pp. 4197–4212, 2021.

[62] A. Alotaibi and A. Barnawi, “Securing massive IoT in 6G: Recent solutions, architectures, future directions,” *Internet Things*, p. 100715, 2023.

[63] D. Sirohi, N. Kumar, P. S. Rana, S. Tanwar, R. Iqbal, and M. Hijjii, “Federated learning for 6G-enabled secure communication systems: a comprehensive survey,” *Artif. Intell. Rev.*, pp. 1–93, 2023.

[64] N. Ebrahimi, H.-S. Kim, and D. Blaauw, “Physical layer secret key generation using joint interference and phase shift keying modulation,” *IEEE Trans. Microw. Theory Tech.*, vol. 69, no. 5, pp. 2673–2685, 2021.

[65] N. Xie, Z. Li, and H. Tan, “A survey of physical-layer authentication in wireless communications,” *IEEE Commun. Surv. Tutor.*, vol. 23, no. 1, pp. 282–310, 2020.

[66] A. Aldweesh, A. Derhab, and A. Z. Emam, “Deep learning approaches for anomaly-based intrusion detection systems: A survey, taxonomy, and open issues,” *Knowl.-Based Syst.*, vol. 189, p. 105124, 2020.

[67] V. Talreja, M. C. Valenti, and N. M. Nasrabadi, “Deep hashing for secure multimodal biometrics,” *IEEE Trans. Inf. Forensics Secur.*, vol. 16, pp. 1306–1321, 2020.

[68] T. Edwards and M. S. Hossain, “Effectiveness of deep learning on serial fusion based biometric systems,” *IEEE Trans. Artif. Intell.*, vol. 2, no. 1, pp. 28–41, 2021.

[69] Y. Al-Eryani and E. Hossain, “The D-OMA method for massive multiple access in 6G: Performance, security, and challenges,” *IEEE Veh. Technol. Mag.*, vol. 14, no. 3, pp. 92–99, 2019.

[70] T. Li, Z. Hong, L. Liu, Z. Wen, and L. Yu, “Meta-WF: Meta-Learning-Based Few-Shot Wireless Impersonation Detection for Wi-Fi Networks,” *IEEE Commun. Lett.*, vol. 25, no. 11, pp. 3585–3589, 2021.

[71] B. Li, Y. Wu, J. Song, R. Lu, T. Li, and L. Zhao, “DeepFed: Federated deep learning for intrusion detection in industrial cyber–physical systems,” *IEEE Trans. Ind. Inform.*, vol. 17, no. 8, pp. 5615–5624, 2020.

[72] I. A. Ridhawi, S. Otoum, and M. Aloqaily, “Decentralized Zero-Trust Framework for Digital Twin-based 6G,” *ArXiv Prepr. ArXiv230203107*, 2023.

[73] M. A. Rahman and M. S. Hossain, “A deep learning assisted software defined security architecture for 6G wireless networks: IIoT perspective,” *IEEE Wirel. Commun.*, vol. 29, no. 2, pp. 52–59, 2022.

[74] E. Rodríguez, B. Otero, and R. Canal, “A survey of machine and deep learning methods for privacy protection in the Internet of Things,” *Sensors*, vol. 23, no. 3, p. 1252, 2023.

[75] I. K. Dutta, B. Ghosh, A. Carlson, M. Totaro, and M. Bayoumi, “Generative adversarial networks in security: a survey,” in *2020 11th IEEE Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON)*, IEEE, 2020, pp. 0399–0405.

[76] E. Rodríguez *et al.*, “Transfer-Learning-Based Intrusion Detection Framework in IoT Networks,” *Sensors*, vol. 22, no. 15, p. 5621, 2022.

[77] A. Clemm, L. Ciavaglia, L. Granville, and J. Tantsura, “RFC 9315 Intent-Based Networking-Concepts and Definitions,” 2022.

[78] “ENISA Transport Threat Landscape — ENISA.” https://www.enisa.europa.eu/publications/enisa-transport-threat-landscape (accessed Jun. 25, 2023).

[79] S. Bradner, “RFC2119: Key words for use in RFCs to Indicate Requirement Levels.” RFC Editor, 1997.

[80] 3GPP, “System architecture for the 5G system,” *3GPP TS 23501 V15 30*, 2018.

[81] 3GPP, “System architecture for the 5G system,” *3GPP TS 23501 V1740*, 2022.

[82] “‘6G Architecture Landscape – European perspective’ White paper.” https://5g-ppp.eu/6g-architecture-landscape-european-perspective-white-paper/ (accessed Jun. 25, 2023).

[83] J. Lin, L. Dang, M. Rahouti, and K. Xiong, “Ml attack models: Adversarial attacks and data poisoning attacks,” *ArXiv Prepr. ArXiv211202797*, 2021.

[84] “Mitigations Against Adversarial Attacks,” *F-Secure Blog*, Jul. 11, 2019. https://blog.f-secure.com/mitigations-against-adversarial-attacks/ (accessed Jun. 25, 2023).

[85] S. Samarakoon *et al.*, “5G-NIDD: A Comprehensive Network Intrusion Detection Dataset Generated over 5G Wireless Network,” *ArXiv Prepr. ArXiv221201298*, 2022.

[86] A. Parmisano, S. Garcia, and M. J. Erquiaga, “A labeled dataset with malicious and benign iot network traffic,” *Stratos. Lab. Praha Czech Repub.*, 2020.

[87] “IDS 2018 | Datasets | Research | Canadian Institute for Cybersecurity | UNB.” https://www.unb.ca/cic/datasets/ids-2018.html (accessed Jun. 25, 2023).

[88] “The Bot-IoT Dataset | UNSW Research.” https://research.unsw.edu.au/projects/bot-iot-dataset (accessed Jun. 25, 2023).

[89] N. Moustafa and J. Slay, “UNSW-NB15: a comprehensive data set for network intrusion detection systems (UNSW-NB15 network data set),” in *2015 military communications and information systems conference (MilCIS)*, IEEE, 2015, pp. 1–6.

# Appendix A

## 8.1. Summary of EU beyond-5G and 6G research projects

**5G-CLARITY** is focused on the development and demonstration of an advanced system that goes beyond 5G to integrate 5G, Wi-Fi, and LiFi technologies within private networks. This system will be managed through AI-based autonomic networking. The project's main objective is to strengthen Europe's leadership in the expanding markets of private 5G networks and 5G for factory automation. By addressing challenges related to spectrum flexibility, delivery of critical services, and autonomic network management, 5G-CLARITY aims to design a comprehensive system for beyond 5G private networks. This system consists of two key pillars: a heterogeneous wireless access network that combines 5G beyond R16, Wi-Fi, and LiFi technologies, and a novel management plane based on Software Defined Networking (SDN) and Network Function Virtualization (NFV) principles. The management plane, powered by AI algorithms, enables network slicing for neutral hosts and facilitates autonomic network management.

**5G-COMPLETE** aims to revolutionize the architecture of 5G networks by efficiently combining compute and storage resource functionalities within a unified ultra-high capacity converged Fiber-Wireless (FiWi) Radio Access Network (RAN). Leveraging advancements in Ethernet fronthauling through the eCPRI standard, 5G-COMPLETE introduces a unique architectural proposition that integrates key technologies. These include the high capacity of fiber and high-frequency radio, converged FiWi fronthauling, spectral efficiency of analog modulation and coding schemes, flexibility of mesh self-organized networks, efficiency of high-speed and time-sensitive packet-switched transport, rapid and cost-efficient service deployment using unikernel technology, and an enhanced security framework based on post-Quantum cryptosystems. The proposed converged infrastructure of 5G-COMPLETE merges the 5G New Radio fronthaul/midhaul/backhaul functionalities into a single Ethernet-based platform, transforming the RAN into a low-power distributed computer and introducing new network concepts. The project's outcomes will be validated through scalable lab and field-trial demonstrators in Athens (Greece), Lannion (France), and Bristol (UK). With the participation of 13 consortium partners spanning the entire 5G research and market chain, 5G-COMPLETE aims to introduce new business models and drive novel research opportunities, ultimately delivering tangible results.

**5GZORRO** aims to develop innovative solutions for zero-touch service, network, and security management in multi-stakeholder environments, utilizing Smart contracts based on Distributed Ledger Technologies (DLT) to enhance business agility. The consortium envisions the evolution of 5G networks to support diverse vertical applications in a highly pervasive shared infrastructure. This will be achieved through automated end-to-end network slicing across multiple operators and resource providers, who can share various types of resources such as spectrum and virtualized access. To enable seamless management and orchestration, 5GZORRO leverages distributed Artificial Intelligence (AI) for cognitive network automation, reducing the need for manual intervention. Additionally, DLT is utilized to implement distributed security and trust mechanisms across the 5G service chain. The overall goal of 5GZORRO is to provide a secure, trusted, and efficient cross-domain solution for zero-touch service and network management, including intelligent resource discovery, brokerage, and workload offloading mechanisms. This encompasses the establishment of a 5G Operational Data-Lake with APIs, a Smart Contract solution for SLA management, cross-domain security and trust mechanisms, and a secure shared spectrum market for real-time trading of spectrum allocations.

**ARIADNE** aims to revolutionize the wireless landscape by shifting the focus from localized network improvements to a vision of pervasive mobile virtual services through an integrated system beyond 5G. By combining a novel high-frequency advanced radio architecture and an Artificial Intelligence-based network processing and management approach, ARIADNE seeks to establish ultra-high spectral efficiency and reliable communications in the bandwidth-rich D-band. This concept will be realized through Machine Learning-based design and intelligent network management, enabling dynamic establishment and reconfiguration of connections in both Line of Sight (LOS) and Non-Line of Sight (NLOS) environments. The project encompasses the development of new radio technologies for communication in the D-Band frequency range, utilization of metasurfaces as tuneable reflectors to shape the propagation environment, and the application of Machine Learning and Artificial Intelligence techniques for effective management and dynamic assignment of metasurfaces. ARIADNE's objectives include experimental validation of theoretical findings, assessment of practical limitations and obstacles, feasibility evaluation of the proposed architecture and components, and identification of accelerators and potential solutions for the adoption of AI-aided D-band wireless in future wireless systems beyond 5G.

The **INSPIRE-5Gplus** project objective is the advancement of security vision for 5G and Beyond devising a smart, trustworthy and liability-aware 5G end-to-end security platform for future connected systems, through the adoption of emerging trends and technologies.

The INSPIRE-5Gplus architecture is designed to provide automated, end-to-end security management for multi-domain 5G environments. It emphasises protection, trustworthiness, and accountability in managing virtualized networks across multi-domains including radio, edge, and core segments. Each Security Management Domain (SMD) is responsible for intelligently managing the security of resources and services within its scope, and takes care of services that span multiple domains, such as end-to-end slicing.

The project has successfully developed a zero-touch security framework that incorporates advanced technologies such as AI techniques, ZSM, TEE, and MTD to meet the needs of 5G use cases. Furthermore, the project validated the HLA through an improved closed loop model focusing on the trust closed loop scenario. Most notably, INSPIRE-5Gplus met its objective and delivered a top-of-the-line security framework with a vast array of features that will support further research and experimentation in the upcoming Beyond 5G networks.

The **LOCUS** project aims to enhance 5G infrastructures by providing accurate and ubiquitous location information and deriving complex features from physical events, making them available to applications via simple interfaces. This enables network operators to manage their networks more effectively and provide new applications and services, thereby increasing the overall value of the 5G ecosystem. The proposed system architecture is an augmentation of the 5G architecture, where network and user data from heterogeneous technologies are combined to extract on-demand analytics, enabling a plethora of new people-centric and network- centric applications for 5G verticals.

The Locus Project has developed a system architecture with built-in security and privacy which utilises 5G Terminal Localization and integrates with non-3GPP localization technologies such as GNSS, WIFi, Bluetooth, etc. The Locus Project has also developed an hybrid virtualization platform that distributes virtual functions across edge and core computing locations within the 5G end- to-end infrastructures combining automation capabilities through LOCUS Management and Orchestration (MANO).

The **MonB5G** Project [Ref1] focus on the development of a zero-touch management and orchestration to support network slicing for 5G and Beyond 5G (B5G) mobile systems. MonB5G architecture consists of a centralized management system subdivided into many management systems, which distribute intelligence and decision-making to different components. The data-driven components are based on a service-oriented architecture (SOA) and use AI techniques and federated learning. The key technologies used in MonB5G architecture include advanced radio technologies, such as Massive MIMO and Beamforming, to improve spectral efficiency; and radio access technologies, such as terahertz communication and visible light communication, to significantly increase the capacity and speed of wireless networks. The main goals of the project are the development of a network optimization tool using AI and ML techniques and the development of a spectrum management platform to enable dynamic and efficient sharing of spectrum resources among different stakeholders.

**TERAWAY** [Ref2] project aims to develop photonic-enabled THz transceivers for 5G wireless communications with high-performance fronthaul (FH) and backhaul (BH) links. TERAWAY integrates optical and photonic technology to provide the foundations for generating, transmitting and detecting radio signals in an ultra-wide range of carrier frequencies (D-band, W-band (W/D) and THz by integrating optical and photonic technologies. The three-fold goal of TERAWAY is to develop photonics-enabled platform for FH and BH link; to improve the performance and energy efficiency, while providing network slice, to different services; and to develop a new SDN controller to manage network and radio resources.

[Ref2] TERAWAY European Research Project, https://ict-teraway.eu/

**6G BRAINS** focuses on addressing current challenges in networks beyond 5G, such as Terahertz (THz) and Optical Wireless Communications (OWC), as well as Artificial Intelligence (AI) and Deep Reinforcement Learning (DRL) techniques. The project aims to unlock improved wireless connectivity by utilizing novel spectrum combinations and applying AI-driven resource management at multiple decision layers. By introducing multi-agent DRL, 6G BRAINS aims to enhance the performance of industrial networks in terms of capacity, reliability, and latency, especially in massive machine-type communications and future industrial scenarios. The project proposes a comprehensive cross-layer DRL-driven resource allocation solution that supports dynamic cell-free networks with the assistance of device-to-device (D2D) communication and high-resolution 3D Simultaneous Localization and Mapping (SLAM). The technologies developed in 6G BRAINS encompass new spectral links, dynamic D2D network modeling, an intelligent end-to-end network architecture integrating multi-agent DRL, and AI-enhanced 3D SLAM data fusion. The solutions will be validated through proof-of-concept trials, with applications spanning various sectors like Industry 4.0, intelligent transportation, and eHealth. The project also aims to identify new business opportunities for future exploitation activities.

**AI@EDGE** aims to overcome the challenges involved in utilizing the concept of secure and reliable AI for network automation. The initiative brings together European industries, academics, and SMEs to achieve a widespread impact on the AI-for-networks and networks-for-AI paradigms in beyond 5G systems. The primary focus areas of AI@EDGE are cooperative perception for vehicular networks, secure and multi-stakeholder AI for IoT, aerial infrastructure inspections, and in-flight entertainment, with the goal of maximizing their commercial, societal, and environmental impact.

To accomplish this, AI@EDGE is concentrating on two significant breakthroughs. The first breakthrough is developing general-purpose frameworks for closed-loop network automation that can support flexible and programmable pipelines to create, utilize, and adapt secure, reusable, and trustworthy AI/ML models. The second breakthrough is building a converged connect-compute platform for creating and managing resilient, elastic, and secure end-to-end slices that can support various AI-enabled network applications.

**DAEMON** aims to take a practical approach to network infrastructure (NI) design, rather than simply following the hype surrounding artificial intelligence (AI). Through a systematic analysis, DAEMON will determine which NI tasks are best suited to be solved with AI models, and provide guidelines for the use of machine learning in network functions. Where AI is appropriate, DAEMON will create customized AI models that address the unique needs of network functions, utilizing the latest advancements in machine learning. Using these models, DAEMON will design an end-to-end NI-native architecture for B5G that fully coordinates NI-assisted functionalities.

The progress made by DAEMON in NI will be applied to real-world network settings in order to achieve several objectives, including delivering high performance while efficiently using radio and computational resources, reducing the energy footprint of mobile networks, and providing increased reliability beyond that of 5G systems. DAEMON will design practical algorithms for eight NI-assisted functionalities, selected to achieve these objectives. The performance of these algorithms will be tested in real-world conditions at four experimental sites, and at scale using data-driven approaches based on two nationwide traffic measurement datasets. Nine ambitious yet achievable key performance indicator (KPI) targets will be used to evaluate the effectiveness of the DAEMON algorithms.

The goal of **DEDICAT** 6G [1] is to create a smart connectivity platform using artificial intelligence and blockchain techniques that will enable 6G networks to combine the existing communication infrastructure with novel distribution of intelligence (data, computation, and storage) at the edge to allow not only flexible, but also energy efficient realisation of the envisaged real-time experience. In this context, DEDICAT-6G, aims to leverage dynamic coverage extension via smart connected IoT devices (e.g., drones, car, robots). Moreover, security and privacy issues are examined as well. The three-fold goal of DEDICAT 6G is to provide an energy efficient 6G infrastructure, reduce capital costs and at the same time minimize latency and response times.

In **Hex-x** various issues related to the proper deployment of 6G networks are addressed, such as AI//ML approaches for holistic network optimization, use of higher spectrum bands, network virtualization and disaggregation concepts [2]. To this end, 6 key technology pillars are identified, and in particular connecting intelligence, network of networks, sustainability, global service coverage, extreme experience as well as trustworthiness.

The **MARSAL** project aims to address the challenges of managing and optimizing network resources in the complex and dynamic ecosystem of 5G and beyond mobile networks. By utilizing a converged optical-wireless infrastructure, the project focuses on developing a comprehensive framework for end-to-end performance analysis and resource orchestration. In the network design domain, MARSAL aims to scale up wireless access points cost-effectively through novel cell-free solutions and contribute to the O-RAN project. Additionally, the project aims to enhance the flexibility of optical access architectures in the fronthaul/midhaul segments. In terms of network and service management, MARSAL incorporates machine learning algorithms to manage communication and computational resources, including edge and midhaul data centers. In the network security domain, the project introduces mechanisms to ensure privacy and security for application workload and data, leveraging artificial intelligence and blockchain technologies to create a secure multi-tenant slicing environment. Overall, MARSAL seeks to enable efficient resource management and deliver a high level of performance in 5G and beyond networks.

The **REINDEER** project aims to develop a scalable smart connect-compute platform with zero latency and the ability to interact with a large number of embedded devices. The project will focus on creating a wireless access infrastructure called "RadioWeaves," which consists of distributed radio, computing, and storage resources functioning as a massive antenna array. Protocols and algorithms will be developed to enable resilient interactive applications for robotized industrial environments, immersive entertainment, and intuitive care. The project will also co-design focusing algorithms and protocols for improved interaction with energy-neutral devices. REINDEER will provide experimental proof-of-concept in versatile testbeds and contribute to European technological leadership, while creating new business opportunities.

Co-funded by the European Union

The vision is to establish a hyper-diverse smart connectivity platform that supports real-time and real-space interactive experiences with efficient energy and bandwidth utilization.

The **RISE-6G** project aims to explore innovative solutions for future 6G networks that transform wireless networks into distributed smart connectivity infrastructure. The project focuses on leveraging Reconfigurable Intelligent Surfaces (RISs), which enable dynamic control of radio wave propagation, to create a wireless environment as a service. The project has three main activities: realistic modeling of RIS-assisted signal propagation, investigating the limits of RIS-empowered wireless communications and sensing, and designing efficient algorithms for networking orchestration. The goal is to create intelligent, sustainable, and programmable wireless environments that offer diverse services beyond the capabilities of 5G.

**TeraFlow** is focused on creating a novel cloud-native SDN controller for beyond-5G networks. This new SDN controller shall be able to integrate with current NFV and MEC frameworks as well as to provide revolutionary features for flow aggregation, management (service layer), network equipment integration (infrastructure layer), and distributed AI/ML-based security evidence for multi-tenancy. The project proposes an integrated solution for tackling various challenges of B5G networks to support service providers and telecommunication operators in their journey towards future networks. The project has led the creation of an ETSI OSG (Open Source Group), an open-source project TFS (TeraFlow SDN Controller) [https://tfs.etsi.org/]

Co-funded by the European Union