Call: H2020-ICT-2020-2

Project reference: 101015956

Project Name:

A flagship for B5G/6G vision and intelligent fabric of technology enablers connecting human, physical, and digital worlds

Hexa-X

Deliverable D1.3

Targets and requirements for 6G - initial E2E architecture

Date of delivery:	28/02/2022	Version:	1.0
Start date of project:	01/01/2021	Duration:	30 months

Document properties:

Document Number:	D1.3
Document Title:	Targets and requirements for 6G - initial E2E architecture
Editor(s):	Bahare Masood Khorsandi (NOG), Marco Hoffmann (NOG), Mikko Uusitalo (NOF), Marie-Helene Hamon (ORA), Björn Richerzhagen (SAG), Giovanna D'Aria (TIM), Azeddine Gati (ORA), Diego Lopez (TID)
Authors:	Bahare Masood Khorsandi (NOG), Marco Hoffmann (NOG), Mikko Uusitalo (NOF), Marie-Helene Hamon (ORA), Björn Richerzhagen (SAG), Giovanna D'Aria (TIM), Azeddine Gati (ORA), Erkki Harjula (OUL), Matti Hämäläinen (OUL), Marja Matinmikko-Blue (OUL), Diego Lopez (TID), Antonio Pastor (TID), Riccardo Bassoli (TUD), Frank H.P. Fitzek (TUD), Kim Schindhelm (SAG), Michael Bahr (SAG), Andreas Wolfgang (QRT), Rafael Puerta (EAB), Pål Frenger (EAB), Hans Schotten (TUK), Bin Han (TUK), Stefan Wänstedt (EAB), Mårten Ericson (EAB), Patrik Rugeland (EAB), Christofer Lindheimer (EAB), Pernilla Bergmark (EAB), Damiano Rapone (TIM), Ignacio Labrador Pavón (ATO), Sławomir Kukliński (ORA-PL), Giada Landi (NXW), Cedric Morin (BCO), Cao-Thanh Phan (BCO), Mehdi Abad (EBY), Merve Saimler (EBY), Elif Ustundag Soykan (EBY), Emrah Tomur (EBY), Peter Schneider (NOG), Ana Galindo-Serrano (ORA), Samuli Vaija, Esteban Selva (ORA), Tommy Svensson (CHA), Panagiotis Demestichas (WIN), Panagiotis Vlacheas (WIN), Ioannis-Prodromos Belikaidis (WIN), Vasiliki Lamprousi (WIN), Serge Bories (CEA), Emilio Calvanese Strinati (CEA), Mattia Merluzzi (CEA), Giacomo Bernini (NXW), Nicola Pio Magnani (TIM), Miltiadis Filippou (INT)
Contractual Date of Delivery:	28/02/2022
Dissemination level:	PU[1]
Status:	Final
Version:	1.0
File Name:	Hexa-X D1.3

[1] PU = Public

Revision History

Revision	Date	Issued by	Description
V0.1	20.07.2021	Hexa-X WP1	First template of deliverable D1.3
V0.2	22/11/2021	Hexa-X WP1	Draft of deliverable D1.3 ready for internal review
V0.3	31/01/2022	Hexa-X WP1	Draft is ready for external review
V0.4	23/02/2022	Hexa-X WP1	Deliverable D1.3 is ready for submission
V1.0	28/02/2022	Hexa-X WP1	Final version of deliverable D1.3

Abstract

This report includes an intermediate status update on use cases, KVIs and spectrum, as well as achievements with respect to sustainability (final targets and progress of the project vs. targets). A draft of the E2E architecture is delivered, including a draft security architecture and an update on security considerations.

Keywords

6G vision, use cases, services, key value indicators, performance, E2E architecture, spectrum, sustainability, security

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101015956.

Executive Summary

This report is the third deliverable of Work Package 1 (WP1) — "Targets and requirements for 6G - initial E2E architecture" — and focuses on 6G Hexa-X novel use cases, services, and Key Value Indicators (KVIs) as well as Key Performance Indicators (KPIs). This report also indicates the impact of sustainability, security, and spectrum evolution aspects for 6G. A first gap analysis for end-to-end (E2E) 6G architecture has also been developed.

In this deliverable, the set of use case families (i.e., groups of use cases gathered according to patterns of common usage, in different activity areas) presented in Hexa-X deliverable D1.2 have been revisited to reach a more consistent set of use cases, reflecting the trends foreseen for 6G. Some **families of use cases** are updated and modified while some new use case families and new use cases are introduced: **enabling sustainability, massive twinning, immersive telepresence for enhanced interactions, from robots to cobots, hyperconnected resilient network infrastructures**. These refinements have been driven by the feedbacks collected on deliverable D1.2 and the envisioned trends for the usage of 6G. A subset of these use cases has been further detailed, with first insights on deployment scenarios.

Updated **KPI** definitions are presented based on qualitative analysis of KPI and KVI areas for a selection of use cases presented in D1.2 as well as gap analysis performed by technical work packages during the first phase of the project. Furthermore, a mapping between KPIs and the **KVI**s related to **trustworthiness**, and **sustainability** is presented. The targeted KPI ranges and the reasoning behind these targets are discussed for selected use cases as well as the relation of the use cases to the United Nation Sustainable Development Goals (UN SDGs) and the KVIs are derived.

This deliverable has the first description of a **6G E2E architecture**. The gap analysis is performed based on use cases defined in D1.2 as well as a set of requirements introduced by each technical enabler. Furthermore, a system view figure of 6G E2E architecture is presented. The figure and the following chapter outline all technical enablers from **RAN technologies and localisation and sensing to enablers for intelligent, flexible, and efficient networks**. The chapter also describes the requirements of **the management and orchestration** from E2E architecture.

An initial 6G security architecture and a description of the six security enablers indicated in D1.2 is presented followed by a discussion on the **concept of level of trust**, relating to the KVI "trustworthiness". An overview of the design principles for **security management and orchestration** is also given.

Spectrum evolution aspects relevant to extending **spectrum utilisation** to new frequency bands to address 6G service requirements and achieve both enhanced and novel applications are addressed, which involve both frequency ranges already in use (i.e., low, mid, and mmw) and **new frequency ranges** (i.e., 100--300 MHz and above). Innovative concepts of **flexible spectrum usage and management** are presented as well.

Sustainability is a key research challenge of Hexa-X. A study is performed on the **E2E vision on environmental aspects,** taking into account, among others**,** the main drivers for future trends in emissions of ICT energy usage and the decarbonisation of ICT energy consumption. Afterward, Hexa-X objectives and targets to be fulfilled are discussed. Last but not least proposals are presented on how 6G can facilitate further sustainability targets.

Table of Contents

List of Figures

List of Tables

List of Acronyms and Abbreviations

4G,5G, 6G	4th, 5th, and 6th mobile network generations
AAS	Active Antenna Systems
ADEME	French Agency for Environment and Energy Management
AES	Advanced Encryption Standard
AF	Application Function
AGV	Automated Guided Vehicle
AI	Artificial Intelligence
AI API	AI Application Programming Interface
AIaaS	AI-as-a-Service
AIS	AI Information Service
AKA	Authentication and Key Agreement
AMF	Access and Mobility Management Function
AMVNO	Autonomic Mobile Virtual Network Operator
AP	Access Point
APE	Abstract Processing Element
API	Application Programming Interface
AR	Augmented Reality
ARPF	Authentication credential Repository and Processing Function
ASF	Abstract Switch Fabric
B5G	Beyond 5th generation
BAN	Body Area Network
C/I	Carrier to Interface
CA	Carrier Aggregation
CaaS	Compute-as-a-Service
CAPEX	Capital Expense
CC	Confidential Computing
CI/CD/CD	Continuous Integration / Continuous Deployment/Continuous Deployment
CK	Ciphering Key
CN	Core Network
CO	Continuum Orchestration
CP	Control Plane
CPU	Central Processing Unit
CSP	Content Service Provider
CVM	Compute Virtual Machine
D2D	Device-to-Device
DAS	Distributed Antenna System
DC	Dual Connectivity
DDoS	Distributed Denial-of-Service attack

DevSecOps	Development, Security and Operation
DFP	Dynamic Function Placement
DL	Downlink
DLT	Distributed Ledger Technologies
D-MIMO	Distributed MIMO
DO	Data Object
DT	Digital Twin
DU	Distributed Unit
E&M	Entertainment & Media
E2E	end-to-end
EaaS	Entropy-as-a-Service
ECC	Elliptic Curve Cryptography
ECIES	Elliptic Curve Integrated Encryption Scheme
EHF	Extremely High Frequency
EIRP	Effective Isotropic Radiated Power
EMG	Electromyography
eSIM	Embedded SIM
FCC	Federal Communications Commission
FL	Federated Learning
FW	Firmware
GDPR	General Data Protection Regulation
GHG	Green House Gas
HAP	High-Altitude Platforms
HetNet	Heterogeneous Network
HMI	Human-Machine Interface
HSM	Hardware Security Module
HW	Hardware
I/N	Interface to Noise
IAB	Integrated Access and Backhaul
IMT	International Mobile Telecommunications
Intra-X	Intra-subnetwork
IoT	Internet of Things
iSIM	Integrated SIM
JRC2LS	Joint Radar, Communication, Computation, Localisation, and Sensing
JT-CoMP	Joint Transmission Coordinated Multi-Point
KMS	Key Management System
KPI	Key Performance Indicator
KVI	Key Value Indicator
LAN	Local Area Network
LCA	Life Cycle Assessment

LCI	Life Cycle Inventory
LCIA	Life Cycle Impact Assessment
LCM	Life-Cycle Management
LEO	Low Earth Orbit
LoS	Line of Sight
LoT	Level of Trust
LSA	Licensed Spectrum Access
M&O	Management and Orchestration
MC	Multi Connectivity
MDAF	Management Data Analytics Function
MEC	Multi-access Edge Computing
MEO	Medium Earth Orbit
MIMO	Multiple Input Multiple Output
ML	Machine Learning
MNO	Mobile Network Operator
MR	Mixed Reality
Multi-TRP	Multiple Transmission and Reception Point
NF	Network Function
NFV	Network Function Virtualisation
NiN	Networks in Network
NN	Neural Network
NPN	Non-Public Network
NR	New Radio
NS	Network Service
NTN	Non-Terrestrial Network
NWDAF	Network Data Analytics Function
OAM	Operations, Administration, and Maintenance
OPEX	Operational Expense
OSS/BSS	Operations Support System/Business Support System
OT	Operational Technology
OTT	Over The Top
PAN	Personal Area Network
PEF	Product Environmental Footprint
PET	Privacy Enhancing Technology
PLS	Physical Layer Security
PPA	Power Purchase Agreement
PPS	Programmable Protocol Stack
PQC	Post Quantum Cryptography
QKD	Quantum Key Distribution
QoE	Quality of Experience

QoS	Quality of Service
QRNG	Quantum Random Number Generators
RAN	Radio Access Network
RF	Radio Frequency
RIS	Reconfigurable Intelligent Surface
RSA	Rivest-Shamir-Adleman encryption scheme
RTT	Round Trip Time
RU	Radio Unit
RVM	Radio Virtual Machine
SBA	Service-Based Architecture
SDN	Software-Defined Networking
SLA	Service Level Agreement
SMF	Session Management Function
SSLA	Secure Service Level Agreement
SUCI	Subscription Concealed Identifier
SW	Software
TCO	Total Cost of Ownership
TEE	Trusted Execution Environment
TPM	Trusted Platform Modules
TSN	Time Sensitive Network
UAV	Unmanned Aerial Vehicle
UDM	Unified Data Management
UDR	Unified Data Repository
UE	User Equipment
UL	Uplink
UN SDG	United Nation Sustainable Development Goals
UP	User Plane
UPF	User Plane Function
URLLC	Ultra-Reliable Low Latency Communication
USIM	Universal Subscriber Identity Module
V2X	Vehicle-to-Everything
VR	Virtual Reality
VSF	Virtual Security Function
WNOS	Wireless Network Operating System
X2I	Subnetwork-to-wide-area network
X2X	Inter-subnetwork
XAI	Explainable AI
XR	Extended Reality
ZSM	Zero-touch network and Service Management

1 Introduction

Hexa-X is one of the 5G-PPP projects under the EU Horizon 2020 framework. It is a flagship project that develops a 6G vision and an intelligent fabric of technology enablers connecting human, physical and digital worlds.

This report is the third deliverable of Work Package 1 (WP1) — "End-to-End Vision, Architecture and System Aspects". It is a superset of the last two deliverables D1.1 — "6G Vision, use cases and key societal values" — and D1.2 — "Expanded 6G vision, use cases and societal values including aspects of sustainability, security and spectrum" — and presents an intermediate status update on use cases and KVIs. This report includes a draft of the requirements and impacts of technical enablers on the E2E architecture. This is the first indication of an end-to-end (E2E) architecture in this series of deliverables. A more comprehensive deep dive into the proposed 6G architecture will be reported in the next deliverable D1.4. The report includes also a draft 6G security architecture and further considerations on the security enablers introduced in the past deliverable. The deliverable also reports on the sustainability targets and objectives of the Hexa-X project as well as provides an E2E vision on the environmental aspects of different sectors. Last, it provides an overview of the spectrum that could be used for 6G, for the low, mid, and mmw (millimetre wave) bands as well as the potential extension into the 100 GHz to 300 GHz frequency range.

1.1 Objective of the document

The objective of this document is to analyse introduced use cases, services, and Key Value Indicators (KVIs) / Key Performance Indicators (KPIs) and to provide high-level requirements and early definitions of the deployment scenarios. An intermediate consideration of sustainability targets as well as spectrum studies is included. This document also provides the first indication of the requirements for an E2E 6G architecture including a security architecture.

The document guides the work in the project (especially of all technology-enabling Work Packages, WP2 - 7) and will be disseminated globally.

1.2 Project and Work Package 1 set-up

The Hexa-X project is structured in nine work packages (see Figure 1-1) spanning a timeframe of 30 months. WP1 — "End-to-End Vision, Architecture and System aspects" — interacts with all the other technical WPs (WP2 – WP7), steering their work and including the research results into a common 6G Hexa-X E2E view. The technical work packages are focused on the design and evaluation of technical enablers and components for B5G/6G. WP8 and WP9 cover horizontal activities related to impact creation and project management, respectively.

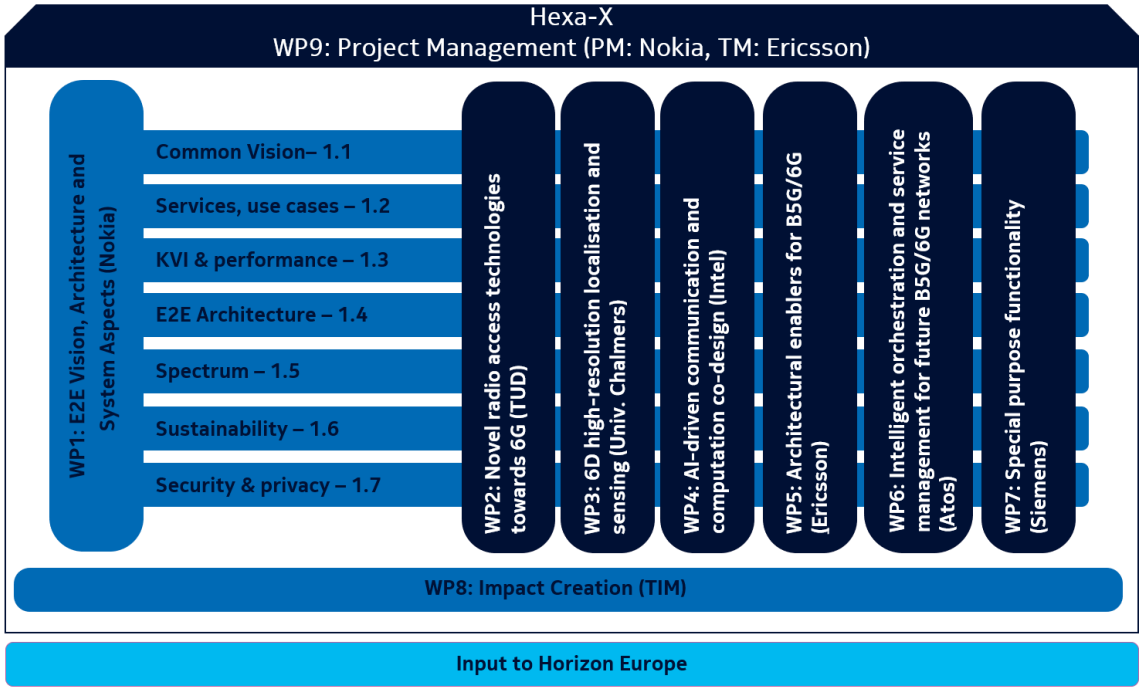

Figure 1-1: ***Hexa-X Project Structure***

1.2.1 Structure and main objective of WP1

This report is Deliverable D1.3 of WP1. WP1 has the main objective to define an overall vision, use cases, and architecture of the x-enabler fabric capable of integrating the technology themes of research connected intelligence, sustainability, trustworthiness, inclusion, and extreme experience. WP1 will guide the work in the whole project and will provide requirements for all other WPs. It covers relevant E2E topics, such as architecture, security, spectrum, KVIs, and KPIs. WP1 is split into seven tasks (Task 1.1 – Task 1.7, see Figure 1-1) to achieve its main goal.

1.2.2 Work plan and deliverables

The set of foundational elements on vision (Task 1.1), use cases and services (Task 1.2), KVIs/KPIs (Task 1.3), E2E architecture (Task 1.4), spectrum (Task 1.5), sustainability (Task 1.6), and security (Task 1.7) have been integrated to help build a seamless and cohesive 6G Hexa-X vision and architecture.

WP1 will provide the following deliverables:

- D1.1: 6G vision, use cases and key societal values (delivered: M02, month two after the project start). This report describes the initial 6G Hexa-X vision including first use cases and KVI aspects.
- D1.2: Expanded 6G vision, use cases and societal values – including aspects of sustainability, security, and spectrum (delivered: M04). This report describes the vision to guide the future research towards 6G. Use cases and KVIs will be identified, providing high-level requirements and definitions of the deployment scenarios. Initial consideration of sustainability targets, spectrum, and security aspects will also be included.
- D1.3: Target and requirements for 6G – initial E2E architecture (delivered: M14). This report includes an intermediate status update on use cases, KVIs, and spectrum, as well as achievements with respect to sustainability (final targets and progress of the project

vs. targets). A draft of the E2E architecture is delivered, including a draft security architecture and an update on security considerations.

- D1.4: Hexa-X architecture for 6G networks – final release (delivery date: M30). This report will present the final Hexa-X vision for 6G, with final use cases, links to the technical work on enablers, final results on sustainability, the E2E Hexa-X architecture including the security architecture and related security guidelines. The document will be disseminated globally to support global discussion on 6G.

All WP1 deliverables are public.

1.3 Structure of the document

The document is structured in the following way: Chapter 2 introduces a selection of use cases and their requirements and performance targets. Chapter 3 describes an initial E2E 6G architecture. Chapter 4 presents security and trustworthiness topics. Chapter 5 focuses on Spectrum evolution aspects and Chapter 6 describes sustainability topics for 6G. The document concludes with the description of the planned next steps in Chapter 7. Annex A is dedicated to a collection of the technical terminologies used in this deliverable as well as all published deliverables by the Hexa-X project.

2 Selection of use cases and their respective requirements and performance targets

Deliverable D1.2 [HEX21-D12] presented the first set of use cases envisioned for 6G, clustered into five use case families (Figure 2-1). D1.2 also presented a set of Key Performance Indicators (KPIs) and Key Value Indicators (KVIs). This deliverable D1.3 will go one step further, updating the set of use cases, based on the feedback received and the latest trends in the ecosystem. A representative subset of use cases is selected, and the associated KPIs and KVIs are detailed.

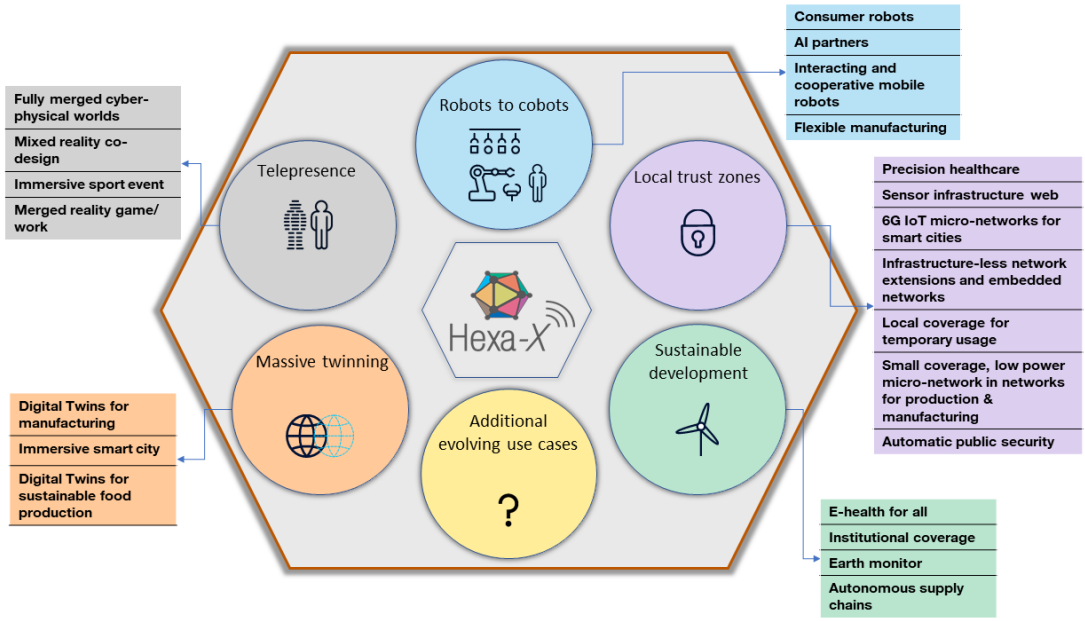

Figure 2-1: ***Hexa-X D1.2 use case families and use cases***

2.1 Updates on use cases

Deliverable D1.2 [HEX21-D12] presented a set of 23 use cases, clustered into 5 use case families, envisioned for 6G in Hexa-X. This set of use cases has been considered as a starting point for the studies in the project. Considering the feedback received after the dissemination and presentation of these use cases inside the project, to the Advisory Group, and in external events, the use cases have been updated by the consortium, to reach a more consistent set, reflecting the trends foreseen in the 6G era. This document only describes the updates with respect to deliverable D1.2: modification of the scope of some use cases, the introduction of a new use case family, and some new use cases. The other use cases are unchanged and their description in D1.2 is still valid. It should be noted that the use cases are included in the most relevant use case family, but this attachment to a use case family is not exclusive as some use cases may have connections to multiple use case families.

2.1.1 “Enabling sustainability" use case family

- The use case family, entitled “Sustainable development” in D1.2, is now renamed as “Enabling sustainability”, to better represent the intention of the use cases included: use of 6G in other sectors than ICT, in order to globally improve sustainability, whether it is about delivering key digital services (such as health services, education, …) in remote or isolated areas or empowering other activity sectors to foster their positive impacts for the planet and society (e.g. reducing waste in production or guaranteeing food production).

This use cases family, therefore, gathers various use cases relying on 6G as a tool to enable sustainability in other activity sectors. Here sustainability is considered at a general level, encompassing environmental sustainability, but also societal sustainability, targeting use cases addressing some of the 17 United Nations Sustainability Development Goals[2]. The use cases in this use case family address different aspects of sustainability, offering to other sectors to leverage on 6G so these new use cases become more sustainable. For example, some of the use cases in this use case family are focused on more efficient usage of resources (such as "autonomous supply chain") while other use cases target protection of the environment ("Earth monitor") or inclusion, with access to crucial services everywhere, even in remote areas (e.g., health services with "E-health for all", or education with "institutional coverage"). The use case labelled "Digital Twins for sustainable food production" in D1.2, part of the "Massive twinning" use case family, has been moved to the "Enabling sustainability" use case family. It has been renamed "Sustainable food production" as this use case illustrates how agriculture can benefit from the opportunities enabled by 6G to better monitor and forecast threats on the food production, while guaranteeing and optimising the production yield. It can, therefore, be seen as an alternative to chemical products, while guaranteeing food production and avoiding shortages. This use case is, therefore, a powerful example of how 6G can contribute to "Enabling sustainability".

- Deliverable D1.2 included a service "Energy-optimised services" in the category "Enabling services harnessing new capabilities", building on a set of global indicators, accounting for the E2E energy consumption (including networks, application, devices) and environmental impact (including material, from construction to recycling) to show/make transparent the possible trade-off between service performance and quality from one side, and sustainability on the other side. It is now moved into the "Enabling sustainability" use case family, as a use case "Network trade-offs for minimised environmental impact" and is illustrated through an example. In the case of the delivery of video services, the user can select from different classes of services, with different trade-offs between the video resolution and E2E environmental cost: ranging from the highest resolution possible for videos, but with high impact on energy consumption and need for new terminals, to the case of video with a degraded, but still acceptable resolution, and minimised energy consumption. Depending on the preferences set by the end-user, they can enjoy the highest video quality and user experiences for videos that they deem worthy and can accept for videos with lower interest a lower level of experience: lower resolution video, higher latency for the download due to a different path selected based on a lower environmental impact. E2E environmental impact should be considered, including the UE. More general, availability of standardised indicators for the energy consumption and footprint are necessary for each part of the network, to enable the tuning of the network (selection of appropriate paths, network management, taking into account criteria such as selecting paths powered by renewable electricity or avoiding paths involving non-circular material usage) towards the targeted trade-off between Quality of Experience (QoE) and environmental footprint.
- A new use case has been introduced, "Network functionality for crisis resilience". As mobile networks are used for more and more fundamental services, the requirement for resilience is amplified especially in extreme situations and crises. Oftentimes, a critical aspect of the network operations is the power supply, which can easily be disrupted during a crisis, e.g., storms, earthquakes, or floods. When this happens, backup power generation

[2] https://sdgs.un.org/

typically kicks in, which oftentimes is realised using small diesel generators. These generators are both less efficient, and more polluting than large-scale power production, as well as providing a finite power source before it needs refuelling. Therefore, it would be very beneficial to minimise the energy consumption, while the network is operating on backup power to extend its operation, while ensuring a sufficient service level. This could entail rerouting traffic through nodes which are still powered through the power grid, when possible, or reducing power-consuming functionality, such as dual connectivity or spatial mapping, and to focus on core functionality, such as providing communication or prioritising first responders' internal calls. It could also entail sacrificing latency performance, especially from idle mode, by increasing the periodicity of the always-on signals (e.g., system information) to reduce the baseline energy consumption. This could also extend to the wireless devices, which, since they are battery-powered, require frequent recharging, typically done using the same emergency backup power (i.e., diesel generators) in case of a crisis. The network could instruct the wireless devices to save power by disabling certain functionality (such as high bandwidth services or processing heavy applications).

2.1.2 "Massive twinning" use case family

- The "Digital Twins for sustainable food production" use case has been moved to the "Enabling sustainability" use case family and renamed as "Sustainable food production".
- Deliverable D1.2 included a service "Internet-of-tags" in the category "Enabling services harnessing new capabilities", moved now into the "Massive twinning" use case family, to illustrate a new use case. This use case generalises and extends the use of tags and the concept of energy harvesting, relying on multiple possible sources (Radio Frequency (RF) waves, solar, …). The massive deployment of tags will enable improved monitoring of the environment, tracking of goods and merchandise, and will allow optimising the different flows. This precise monitoring of assets will allow tracking of the stock, improve the efficiency of the supply chain, and will be a tool to improve circularity, minimising waste, and improving reuse.

2.1.3 "From robots to cobots" use case family

The use case family has been extended with two use cases:

- Deliverable D1.2 included a service "Flexible device type change" in the category "Enabling services harnessing new capabilities", moved now into the "From robots to cobots" use case family, to illustrate a new use case. This use case, entitled "Situation-aware device reconfiguration" still builds upon the capability of a device to detect the environment (e.g., industrial, automotive) and, based on such awareness, to self-adjust to communication needs of the involved service by repurposing its function. Such device operation adjustment includes modifications of device settings and reconfiguration needed to satisfy application performance requirements different from the ones considered in another environment the device used to operate before. This use case will be all the more relevant with the generalisation of robots, achieving a wide variety of tasks, and will allow a robot to perform multiple different tasks, beyond the original task the robot was intended to for a given operating environment. An example could be an industrial robot toggling between critical and non-critical actions, such as switching from welding, requiring very high localisation accuracy and low latencies, to long-range movement only requiring moderate localisation accuracy and latencies.

2.1.4 "Hyperconnected resilient network infrastructures" use case family

The use case family "Local trust zones for human and machines" described in D1.2 is now divided into two new use case families: "Trusted embedded networks" and "Hyperconnected resilient network infrastructures". "Hyperconnected resilient network infrastructures" use case family gathers the use cases involving sub-networks, or networks of networks, requiring a high level of resilience:

- "Sensor infrastructure web" use case (as described in D1.2).
- Deliverable D1.2 included a service "AI-assisted Vehicle-to-Everything (V2X)" in the category "Enabling services harnessing new capabilities", moved now into the "Hyperconnected resilient network infrastructures" use case family. The idea is to exploit Artificial Intelligence (AI) algorithms for enhanced automotive services provided by future 6G networks. More specifically, AI algorithms will be used to process big and raw data gathered not only by sensors (in and outside of the cars) but also by radio stations in the operators' networks in order to perform monitoring and dynamic shaping of the traffic flow as well as to suggest actions/recommendations to connected vehicles' drivers. AI algorithms can also potentially be used to directly control automated vehicles – so to reduce the car traffic caused by them and improve safety for drivers and passengers. This use case has commonalities with the "Sensor infrastructure web" use case in how raw data is gathered from sensors: basically, by means of AI, a real-time digital replica of the real traffic scenario within an entire urban area will be created and dynamically adapted, with the aim to improve safety and mobility's sustainability in a crowded urban environment. What differs with respect to "Sensor infrastructure web" is the subsequent provisioning of actions/recommendations (derived from the AI-based processing of data) to assist the drivers while moving within the concerned urban area. As an example of sustainable mobility, a vehicle could provide a Mobile Network Operator (MNO)-controlled 6G network with information regarding e.g., its own location within a certain urban area for which a digital replica of the traffic scenario is developed. Then, to support such vehicle's smooth mobility, the most appropriate network node, e.g., the roadside unit or the operator's base station, collects and elaborates all the information provided by a multitude of other vehicles as well as other relevant actors (e.g., pedestrians) and entities (e.g., infrastructure monitoring, traffic lights, etc.), all located within a portion of the urban area under the network node's service coverage. Such network node may be chosen e.g., by considering the workload of the node itself along with its own AI processing capabilities. The network nodes within the concerned urban area coordinate with each other by sharing their own AI-processed data so as to define actions/recommendations to be then provided to the vehicle in order to dynamically control and shape the traffic in the whole urban area aiming at reducing the traffic itself.
- "Interconnected IoT micro-networks" use case, formerly described in D1.2 as "6G IoT micro-networks for smart cities", generalised here beyond smart cities, which could benefit from these micro-networks (e.g., agriculture).
- "Enhanced public protection" use case. Often security events are handled unsatisfactorily (e.g., in terms of speed of handling, impacted citizen localisation, general impact mitigation, etc.) due to the lack of information that leads to situation awareness. 6G has a role to play here, through various fundamental blocks, envisaged and developed in Hexa-X. E.g., the "network of networks" approach can lead to a network federation for connecting responders, and other stakeholders that need information. Likewise, advanced radio and management solutions can offer the necessary capacity, even in the event of faults. Finally, and most importantly, as protection is coupled with aspects like

monitoring, tracking, etc., the work on AI-governance, also envisaged for 6G infrastructures, as well as privacy-preserving techniques, secure identity management, and leak-free networks, has a fundamental role to play: to ensure the proper, trustworthy, ethical application of these solutions, i.e., limited to the time, place, and context in which they are needed.

2.1.5 "Trusted embedded networks" use case family

The use case family "Local trust zones for human and machines" described in D1.2 is now divided into two new use case families: "Trusted embedded networks" and "Hyperconnected resilient infrastructures". "Trusted embedded networks" use case family gathers the use cases involving sub-networks, or networks of networks, requiring a high level of trustworthiness:

- "Human-centric communications" formerly described in D1.2 as "Precision healthcare", where the use of in-body and on-body sensors is generalised beyond health purposes.
- "Infrastructure-less network extensions and embedded networks" use case (as described in D1.2).
- "Local coverage for temporary usage" use case (as described in D1.2).
- "Small coverage, low power micro-network in networks for production & manufacturing" use case, as described in D1.2.

2.1.6 Updates on services harnessing new capabilities

Due to the movements described in previous sections, the "Enabling services harnessing new capabilities" category now focuses only on services building on new capabilities, offered to the network "as a service":

- Compute-as-a-service (CaaS[3]), as described in D1.2.
- AI-as-a-service (AIaaS), as described in D1.2
- Security-as-a-service for other networks. As described in D1.2, there is a series of different security features suitable to be considered, specifically related to trusted local connections, on-demand deployment of security functions, and security of E2E paths by the composition of trusted segments. The application of the identified security enablers to address these features requires further investigation.
- Sensing-as-a-service is a new service that will enhance and extend 6G. From weather condition monitoring, to detection and tracking of objects not carrying UEs, to material sensing, are some of the many conceivable services conceivable. These new methods and features will impact next generation 6G architecture in many ways and must be reflected in the service-based architecture. New services and new interfaces for sensing capabilities must be defined as well as existing localisation services extended and potentially revised.
- Positioning-as-a-service is another new service extending the 6G service offering with versatile positioning capabilities. Examples are an extension from 3D to 6D (3D position and 3D orientation) location information as well as high accuracy and low latency provisioning (duration between the initialisation of localisation procedure and acquiring localisation estimate).

[3] CaaS is also used in the literature to represent "Container-as-a-Service", but in Hexa-X, CaaS refers to "Compute-as-a-Service", as introduced in deliverable.

2.1.7 Summary: updated view of Hexa-X use cases

The figure below presents the updated view of Hexa-X use cases and use cases families, highlighting with a blue text the updates compared to the first set of use cases and use case families presented in D1.2.

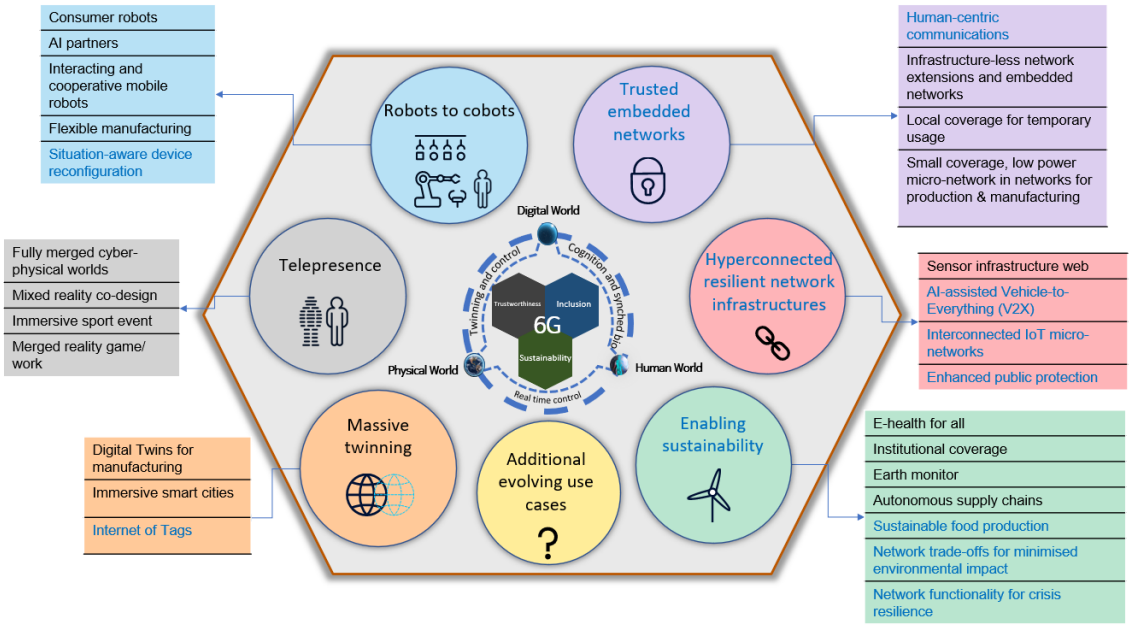

Figure 2-2*: Hexa-X updated use cases*

2.2 Updates on KPI/KVI methodology and definitions

Deliverable [HEX21-D12] presented a qualitative analysis of KPI and KVI areas for a selection of use cases and formulated functional requirements for the respective use cases. Based on this analysis, technical work packages identified and refined KPIs related to the service offered by the work package (e.g., communication, AI, localisation, sensing, ...) in their gap analysis during the first phase of the project. These individual inputs on KPIs have been aligned in WP1 and, consequently, updated KPI definitions are presented in the following in Section 2.2.1.

Following the methodology introduced in [HEX21-D12], some of these KPIs act as proxies for the quantification of KVIs related to, e.g., trustworthiness, flexibility, or sustainability. These are discussed in more details in Section 2.2.2.

The targeted KPI ranges and the reasoning behind these targets (including pointers to the state of the art and technical work in other Hexa-X work packages) are discussed in Section 2.3 for the selected individual use cases. In addition, the relation of each use case to the United Nations Sustainable Development Goals[4] and the Hexa-X key values is described, and indicators are derived.

[4] https://sdgs.un.org/

2.2.1 KPI definitions

KPIs are grouped by services utilised within a use case. These services are *communication*, *AI/computation*, and *localisation and sensing*. The intention behind specifying KPIs for each service individually is maintaining degrees of freedom regarding the realization (e.g., the architecture, the deployment, and orchestration aspects) at this stage. Therefore, the specification of target numbers for these KPIs need to be accompanied by a deployment description, as provided for our selected use cases in Section 2.3. This also implies that aspects such as scalability or resource constraints are not specified as target KPIs but instead as part of the deployment characteristics.

To illustrate this, consider the example of AIaaS being utilised within a use case: from an application or use case perspective, a maximum service Round Trip Time (RTT) is specified as the delay that can be tolerated between issuing a request to the service and the availability of the response. However, depending on the deployment and, thus, the realisation of the architecture in a concrete case, this RTT might include additional communication delays, which are further impacted by orchestration choices. By stating the service RTT from an application perspective (i.e., E2E), the overall time budget can be realised by means of flexible orchestration or resource allocation: e.g., reducing computation time by allocating more powerful resources potentially at the expense of increased communication delay, versus local, less powerful computational resources at reduced communication delays.

To derive more specific KPIs for, e.g., the air interface, the deployment setup, and KPI targets need to be mapped to the architecture to identify communication paths and involved entities that consume the overall latency (or, more general: resource) budget.

Table 2-1 summarizes the KPIs and their definitions. More detailed descriptions are provided in the respective references, both from the available state-of-the-art and other Hexa-X deliverables. For each service, KPIs are grouped into *dependability attributes* characterizing the application demands for the respective service and the underlying *Quality of Service (QoS) attributes* that need to be fulfilled to achieve service availability according to application demands or desired Quality of Experience (QoE).

Table 2-1: *KPI definitions for services*

		KPI	**Definition / References**
Communication	Dependability Attributes	Availability [%]	Percentage of time during which QoS targets are met and service is offered during operation [22.261]. Additional information on the resilience of the application, such as survival time [ms] as acceptable downtime of a service, should be included.
		Reliability [%]	Percentage of the amount of sent packets delivered within QoS constraints [22.261].
		Safety	Reference to applicable standards and regulations in the domain; functional requirements, if available.
		Integrity	Reference to applicable standards and regulations in the domain; functional requirements, if available.
		Maintainability	Functional requirements and regulations (e.g., time to recovery, auditability).
	QoS Attributes	Service latency [ms]	E2E latency for communication service between two application endpoints from use case perspective with allowed variability/Jitter and/or expected upper bound. One should distinguish between Up Link (UL) and Down Link (DL) and detail expected packet sizes

			and traffic characteristics in the deployment characteristics, if available.
		Data rate (minimum expected, desired, maximum) [Mbit/s]	Referring to a single UE. Minimum expected data rate ensures correct operation of the use case (with reduced QoE or limited functionality). Desired data rate ensures desired QoE and full functionality of the use case. If an upper bound can be specified, this is the maximum data rate that can occur in the use case. [5GP21]
		Resource constraints	*Refer to deployment description (e.g., frequency, energy consumption)*
		Scalability	*Refer to deployment description (e.g., number of users, mobility, ...)*
AI and computation	Dependability Attributes	Agent availability [%]	Percentage of time during which the AI agent can receive and respond to an inferencing request (i.e., the agent can be utilised for decision making) meeting the agreed QoS targets.
		Agent reliability [%]	Percentage of requests that are fulfilled within the agreed QoS targets.
		Safety	Requirements or regulations related to the use of AI in the application domain.
		Integrity	Requirements or regulations related to guaranteeing that AI/compute operates as intended.
		Maintainability	Functional requirements or regulations as to how the system reacts to defects/faulty operation (e.g., how well and fast the AI agent recovers from attacks).
	QoS Attributes	AI service RTT [ms]	Maximum tolerable delay from the request being issued by the application until response available to the application by the AI service.
		Inferencing accuracy [%]	Domain-specific measures for the accuracy of the AI estimates, if available (i.e., quality functions).
		Interpretability level	Qualitative indicator for the interpretability of AI-based models and decisions, use case-specific.
		Training/model transfer latency [ms]	Tolerable time to train a new/evolved model relevant in the use case or transfer a model from one AI agent to another AI agent.
		Resource constraints	*Refer to deployment description (e.g., frequency, energy consumption)*
		Scalability	*Refer to deployment description (e.g., number of users, mobility, ...)*
Localisation and Sensing	Dependability Attributes	Service availability [%]	Percentage of time during which location or sensing service requests are answered and the given QoS targets are fulfilled.
		Service reliability [%]	Percentage of requests that are fulfilled within the agreed QoS targets.
		Safety	Requirements or regulations for localisation or sensing regarding safety in the respective use case.
		Integrity	Requirements or regulations for localisation or sensing integrity, e.g., robustness to potential disturbances or attacks.

		Maintainability	Functional requirements or regulations in the use case domain when it comes to utilisation of the service.
	QoS Attributes	Location accuracy [m]	Accuracy of the estimated location, reported in horizontal and vertical position accuracy [22.071].
		Localisation/Sensing service RTT [ms]	Maximum tolerable service latency from the request being issued by the consumer (application, service) until location/sensing response being provided. Not to be confused with RTT measurements as one method to measure the distance between user and base station.
		Orientation accuracy [°]	Accuracy of the estimated direction of UE: roll, pitch, yaw.
		Refresh rate [1/s]	The rate at which new location estimates need to be obtained by the application.
		Minimum and maximum resolvable range [m]	Required minimum and maximum distinguishable distance between two objects for the application/use case.
		Angular resolution [°]	Required minimum distinguishable angle between two objects (orientation).
		Minimum and maximum velocity [m/s]	Velocity range (minimum and maximum velocity) of the object that needs to be measured by the service.
		Velocity resolution [m/s]	Minimum measurable change in velocity of the object.
		Resource constraints	*Refer to deployment description (e.g., frequency, energy consumption)*
		Scalability	*Refer to deployment description (e.g., number of users, mobility, ...)*

2.2.2 Quantification of key values, KVIs

Following the methodology outlined in [HEX21-D12], we discuss the contributions and relation to the UN SDGs for each of the selected use cases in the following section. This corresponds to the "6G for sustainability" Hexa-X value, viewing 6G as an enabler for use cases that increase or support sustainable development in other sectors. In addition to the use cases outlined in the following section, more details on "earth monitor" and "autonomous supply chains" with respect to their potential to increase sustainability in other sectors are provided in [HEX21-D71]. The "sustainable 6G" perspective and potential KPIs for the quantification of sustainability in the 6G system is further discussed in Section 6.

Trustworthiness and a "level of trust" as KVI is discussed in more detail in Section 4.2. The dependability attributes of a service should correspond to user intent when consuming the service and their fulfilment may, therefore, increase the level of trust of the user experiencing the service. Service-specific KPIs such as the interpretability level of AI algorithms can further contribute to the trustworthiness from a user's perspective.

Fulfilment of scalability requirements as outlined in the deployment descriptions for each of the use cases discussed in the following serves as a proxy for the key value "flexibility". In addition, the ranges provided for, e.g., velocity or resolvable range supported by a localisation and sensing service further serves as an indicator for the flexibility of the service. For AI, a corresponding KPI is the training/model transfer latency as defined in the previous section.

2.3 Description and performance targets for selected use cases

In the following, we detail deployment characteristics, KPI targets, and relation to KVIs for the set of selected use cases.

2.3.1 E-Health for all

How the scenario of "E-Health for all" is defined has a major impact on target KVIs and KPIs. Consequently, deployment characteristics and KPIs are outlined for specific aspects of the overall use case in the following. If the goal is to reach 100% of the people around the globe, technical requirements are much more challenging than the ones touching only 99% or less of the population. The last percentage of the population is located in the periphery, very rural places, far away from any cities and villages. To reach those people requires the use of heterogeneous connectivity technologies. Moreover, remote surgery or imaging versus remote monitoring have completely different requirements for system specifications. Data, latency, or computational requirements have the highest priority. Moreover, E-Health provision in the 6G era has also counterparts, e.g., in the digital twins or robotics domains, even in the infrastructure-less areas lacking infrastructure. Thus E-Health can be seen as a cross-discipline solution, which is adopting various KPIs and KVIs.

2.3.1.1 Deployment Characteristics

Table 2-2: ***Characteristics for E-health for all***

Environment	Outdoor: Urban / suburban / rural / anywhere Indoor: hospitals, health centers, nursing homes, care homes, temporary care points, homes, ambulances
Type of deployment	Fixed deployment at hospitals, care centres, nursing homes, care homes, and mobile deployment in ambulances. Temporary deployments at temporary care points, homes, outdoors. Fully controlled environment at hospitals, health centres, nursing homes, care homes. Semi-controlled at temporary care points, homes, ambulances, outdoors. High data rate for video monitoring and real-time video/other parameter monitoring indoors/outdoors. Long-range Internet of Things (IoT) for wireless health monitoring outdoors. Short-range IoT (Personal Area Network (PAN) & Body Area Network (BAN)) for wireless health monitoring indoors.
Users / devices	Users: healthcare professionals, patients/clients, administrators Devices: smart medical instruments (SpO_2/blood pressure monitors, oxygen masks, insulin pumps, etc.), desktop and laptop PCs, tablets, smartphones, wearable devices (smart watches/rings/clothes, etc.), on-body sensors/actuators (smart plasters, stickers, etc.), in-body sensors/actuators (implants, pills, injectables, etc.)

	Number of users/devices: typically, one patient in a session with one-to-many different devices. In interactive scenarios, the typical setup is a one-to-one discussion between a patient and a healthcare professional. The number of concurrent sessions depends on the use case, ranging from a few sessions to thousands of sessions. Connection density varies from very dense installations to sparse deployments.
Mobility	In a hospital, health centre, nursing and care home, and home scenarios, typically static or limited mobility with walking speed inside buildings. In outdoor and ambulance scenarios, the degree of mobility can be high, and velocity can vary heavily from walking to driving speed.
Frequency bands	mmWave or THz communications for data-rate intensive, latency-sensitive and high accuracy applications. Sub-GHz as well as sub-6 GHz or satellite bands for providing coverage at home, outdoors, and ambulance scenarios for various use-cases. Requires licenced and non-licenced bands (if compatible with latency and reliability requirements in the latter case).
Environmental constraints	No known constraints. Specific regulations might apply for some use cases.
Any other constraints	High level of reliability, security, privacy, and trust in any E-Health use case, due to sensitive nature of data and mission-criticality of more advanced use cases. Real-time constraints in interactive and critical monitoring use cases.

2.3.1.2 KPI targets

Table 2-3: ***Targets KPIs for E-Health for all***

		KPI	Target value	Reasoning / References
Communication	Dependability Attributes	Availability [%]	99.99 – 99.9999	High or ultra-high Emergency The key issue is that the user should expect to have a ubiquitous network access anywhere [ETS20a].
		Reliability [%]	99.999 – 99.999999	High or ultra-high Surgery Imaging Emergency
		Safety		Industry-specific regulations apply for human-machine interaction.

		Integrity		Integrity protection and protection against 3rd party usage/modification of data.
		Maintainability		No use-case-specific requirements are known.
	QoS Attributes	Service latency [ms]	0.1 - 100	Depending on the service. Lower latency is required in indoor, robotic-assisted surgery operations. More relaxed requirements in rural, outdoor connections.
		Data rate (minimum expected, desired, maximum) [Mbit/s]	100 kbit/s – 25 Mbit/s	From 100 kbps (sensor data) to 25 Mbit/s (4K video) expected bit rates. Peak bit rates can be much higher for specific applications, e.g., XR remote diagnostics.
		Resource constraints		*refer to deployment characteristics (prev. section)*
		Scalability		*refer to deployment characteristics (prev. section)*
AI and computation	Dependability Attributes	Agent availability [%]	N/A	No use-case-specific requirements are known. The expectation is: same as for a communication service. AI services can be provided offline to ease requirements on availability and reliability.
		Agent reliability [%]	99.9999	Expectation. Target value depends on the use of AI in the use case.
		Safety	high	Expectation. Target value depends on the use of AI in the use case.
		Integrity	high	Expectation. Target value depends on the use of AI in the use case.
		Maintainability	high	Expectation. Target value depends on the use of AI in the use case.
	QoS Attributes	AI service RTT [ms]	NA	AI services can be provided offline, utilizing big data analytics. RTT is not an issue. For real-time assistance, latency requirements and additional regulations would apply.
		Inferencing accuracy [%]	99.999	Health-related decisions do not tolerate faults. Requirements might be more stringent for real-time decision-making and assistance functions.
		Interpretability level	high	Increased demands due to liability.
		Training/model transfer latency [ms]	N/A	Training can occur offline, no requirements in the use case.
		Resource constraints		*refer to deployment characteristics (prev. section)*
		Scalability		*refer to deployment characteristics (prev. section)*

Localisation and Sensing	Dependability Attributes	Service availability [%]	99	Depends on the application at hand. For global coverage aspects, availability might be lower than for remote surgery applications.
		Service reliability [%]	99.999	Health data is highly sensitive, and its reliability requirements are very high
		Safety		SAR limits
		Integrity	99.999	Highly sensitive health data needs to be protected and secured against modification.
		Maintainability	N/A	No use-case-specific requirements are known.
	QoS Attributes	Location accuracy [m]	0.001 - 1	Global coverage option is not prone to high accuracy requirements. On the other hand, assisted surgery requires better than sub-mm accuracy.
		L/S service RTT [ms]		Depends on the application. Rural deployment of basic healthcare service provision has looser requirements than in in-hospital or surgery applications.
		Orientation accuracy [°]		High accuracy requirements might apply for, e.g., assisted surgery. No specific requirements in other aspects of the use case.
		Refresh rate [1/s]		Depends on the application. Basic vital sign measurement does not need a high refresh rate. However, surgical robots require higher rates.
		Minimum and maximum resolvable range [m]	0.0001 - 1000	Depends on the application and service.
		Angular resolution [°]		Depends on the application.
		Velocity range [m/s]		Expected to be small. No use-case-specific requirements.
		Velocity resolution [m/s]		Expected to be small. No use-case-specific requirements.
		Resource constraints		*refer to deployment characteristics (prev. section)*
		Scalability		*refer to deployment characteristics (prev. section)*

2.3.1.3 Quantification of key values, KVIs

Enabling E-health for all is connected to improved health but can also be seen as a key step towards digital inclusion. This use-case can be linked to UN SDG 3 "*Ensure healthy lives and promote well-being for all at all ages*" and especially SDG 3.8: "Achieve universal health coverage" and SDG 3C "Increase health financing and support health workforce in developing countries". By enabling new technological solutions, reliable and real-time healthcare services can be provided independently of the users' location. It is also possible to offer specialised

healthcare services to the regions, which are far away from institutional service providers or lack dedicated healthcare specialists. By enabling e-health for all, there is also a possibility to improve the universal access to reproductive health and rights (SDG 5.6) and develop effective, accountable, and transparent institutions at all levels (SDG 16.6). All examples are related to both sustainability and digital inclusion and the latter one is also to trustworthiness. With remote access to health data and doctors, a reduction in transport could be achieved, contributing to SDG 13.2.

2.3.2 Digital Twins for manufacturing

The Digital Twin (DT) is a virtual replica of a physical asset or process that connects to and receives data from the latter. In the context of Industry 4.0 [BMWi19], wide-spread use of digital twin s results in massive twinning and enables a real-time and accurate 4D (space and time) digital representation of the industrial manufacturing environment.

These digital twins need to be dynamic to continuously reflect the state of the asset or process that it is replicating. This requirement imposes stringent latency and reliability constraints on the communication between the sensors and the digital twin. Further, massive twinning also imposes constraints on throughput since several devices need to be mirrored and kept updated simultaneously.

2.3.2.1 Deployment Characteristics

Table 2-4: ***Deployment characteristics for Digital Twins for Manufacturing***

Environment	Semi/fully controlled outdoor/indoor industrial environment depending on the use case. Outdoor environment includes ports and mines, and indoor includes factories.
Type of deployment	Small cell/sensor network over a mix of licensed and unlicensed bands
Users / devices	Predominantly machine-to-machine traffic, with human traffic originating from operators monitoring/maintaining the factory Devices include sensors, alarms, fixed machinery, and moving autonomous vehicles Up to 100 UEs over a service area of 50m x 10m x 10m (length x breadth x height)
Mobility	Mix of static and mobile UEs. Static UEs are sensors or are UEs on fixed machines. Mobile UEs are either mounted on autonomous vehicles with known and controllable location and trajectory or handheld by operators on the factory floor
Frequency bands	Both high and mid/low-frequency bands. mmWave or THz communication for data-rate intensive and latency-sensitive applications. Sub-6 GHz for providing coverage.

Environmental constraints	Challenging propagation environments such as in mines and ports where it is difficult to maintain radio coverage to continuously update the digital twin. In the specific context of a digital twin of a radio environment, propagation environments such as indoor factories or warehouses where the presence of unpredictable clutter and its time-varying nature makes it difficult to create and maintain a digital twin.
Any other constraints	Trustworthiness regarding security and privacy: When digital twin is controlling a manufacturing environment that involves humans, privacy and ethical frameworks are necessary to protect worker rights and prevent individuals from being identified. Scalability: We require that the increase in computational complexity and communication overhead with respect to the number of DTs or sensors maintaining the digital twin be linear or sub-linear (for example, logarithmic).

2.3.2.2 KPI targets

Table 2-5: ***Target KPIs for Digital Twins for Manufacturing***

		KPI	Target value	Reasoning / References
Communication	Dependability Attributes	Availability [%]	99.99 –99.999999	Taken from [22.104]. Higher target value for use-cases such as motion control and lower for use cases such as process and asset monitoring. Exact value depends on the service.
		Reliability [%]	99.9 – 99.999999	Network layer packet reliability [22.261]. Target value depends on the service. Lower reliability for process and asset monitoring and higher reliability for motion control and alarms.
		Safety	Critical	Constraints are set by the underlying use case/application. Critical for use cases involving humans on the factory floor working in close proximity with machines that are controlled by the digital twin.
		Integrity	Critical	Industry-specific norms and regulations apply, e.g., [ISO20] and activities in IEEE, e.g., P2806.1[5].

[5] https://standards.ieee.org/ieee/2806.1/10370/

		Maintainability	High	Low tolerance to downtime in use cases where the digital twin is essential for maximizing the efficiency of the factory.
	QoS Attributes	Service latency [ms]	0.1 – 100	Depending on the service. Tighter latency requirements for use cases such as motion control. Less stringent latency requirements for use cases such as process and asset monitoring. Both values are taken from [22.104]. Tolerable Jitter is taken as 10% of the targeted service latency.
		Data rate (minimum expected, desired, maximum) [Gbit/s]	Peak: 10 – 100 Average: 1 – 10	[HEX21-D71]
		Resource constraints		*refer to deployment characteristics (prev. section)*
		Scalability		*refer to deployment characteristics (prev. section)*
AI and computation	Dependability Attributes	Agent availability [%]	99.99 – 99.999999	Same as communication service availability. We do not expect agent availability and reliability requirements to be more stringent than those placed on the communication link.
		Agent reliability [%]	99.9 – 99.999999	Same as for network layer packet reliability.
		Safety	Critical	If a digital twin is utilised for controlling robots, safety constraints for human-machine co-working need to be guaranteed, e.g., [ISO20].
		Integrity	High	
		Maintainability	High	Low tolerance to downtime in use cases where the digital twin is essential for maximising the efficiency of the factory.
	QoS Attributes	AI service RTT [ms]	-	Not different from the QoS of any conventional algorithm. The computations of an AI model for inferencing are deterministic in time.
		Inferencing accuracy [%]	-	Use-case dependent.
		Interpretability level	high	Especially in safety-related applications (e.g., in control applications), the explainability of utilised algorithms is a key requirement.
		Training/model transfer latency [ms]	-	Assuming offline training.

		Resource constraints		*refer to deployment characteristics (prev. section)*
		Scalability		*refer to deployment characteristics (prev. section)*
Localisation and Sensing	Dependability Attributes	Service availability [%]	99.99	[HEX21-D31]
		Service reliability [%]		No use-case-specific requirements are known.
		Safety	Critical	Constraints are set by the underlying use case/application. Critical for use cases involving both humans and machines on the factory floor.
		Integrity	High	Constraints are set by the underlying use case/application. For robots/ automated guided vehicles (AGVs) operating alongside humans and controlled by a digital twin, location information has to be robust to attacks.
		Maintainability	High	Depends on the service/use case. High when the digital twin is essential for smooth functioning of the factory.
	QoS Attributes	Location accuracy [m]	cm-level	[HEX21-D31]
		L/S service RTT [ms]	10 – 1000	[22.804]
		Orientation accuracy [°]	Yaw: 0.1 – 5 Pitch + roll: 1 - 10	Accuracy of heading (yaw) is more important than pitch and roll in the case of AGVs
		Refresh rate [1/s]	10 to 10000	[HEX21-D31]
		Minimum and maximum resolvable range [m]	cm-level	[HEX21-D31]
		Angular resolution [°]	Sub-degree level	[HEX21-D31]
		Velocity range [m/s]	± 8	[HEX21-D31]
		Velocity resolution [m/s]	0.5	[HEX21-D31]
		Resource constraints		*refer to deployment characteristics (prev. section)*
		Scalability		*refer to deployment characteristics (prev. section)*

2.3.2.3 Quantification of key values, KVIs

The use case *digital twins for manufacturing* is related to SDG 12: "*responsible consumption and production*" and SDG 9: "*industry innovation and infrastructure*." By enabling E2E monitoring and a digital representation of the manufacturing process, inefficiencies can be detected, and processes can be optimised in terms of resource consumption or waste production (fully automated or with human assistance), related to SDG 9.4 and SDG 12.5. Twinning increases the resilience of the production infrastructure, as the digital representation can be utilised to evaluate the impact of disturbances and potential mitigation strategies (SDG 9.1).

The digital twin further contributes to SDG 8: "*decent work and economic growth*". By enabling new ways to analyse and optimise manufacturing processes based on the vast amount of data gathered in the digital representations, higher economic productivity can be achieved (SDG 8.2). In addition, access to digital twin will generate novel business opportunities also for small and medium enterprises (e.g., for predictive maintenance, process optimisation, human-machine interaction, data representation, and visualisation), contributing to SDG 8.3. Allowing interaction through digital twins will lead to a reduction of working in hazardous conditions for specific use cases. This contributes to SDG 8.8.

2.3.3 Fully merged cyber-physical worlds

Mixed Reality (MR) - a term for advanced augmented reality bringing immersive experiences with more than visuals and audio, adapted to the environment you are in - will be the main interface of communication as we enter the new paradigm of lightweight head and body-worn devices, embedded in our clothing, on the body and other novel user interfaces. It can also be referred to as Extended Reality With 6G, the experience will become fully immersive, blurring the frontier between the physical world and virtual world. People will have multiple wearables that seamlessly interact with each other, through natural, intuitive interfaces. The devices and applications will be fully context-aware, and the network will become increasingly sophisticated at predicting our needs.

Just as we have seen daily routines of communication, meetings, work, shopping, and entertainment moving from the physical world into smartphones, we expect all smartphone applications to move into MR. The immersive qualities of MR will also increase the possibility to further digitalise our daily lives. In very densely populated spaces, such as city centres, shopping malls, and train stations, we can expect humans to use communication services to be efficient, get information, stay in touch, or be entertained just like today. Some examples of applications used in this situation are:

- Holographic communication and telepresence - Via holographic telepresence, it will be possible to sense being virtually in a certain location while really being somewhere else. The user would experience the world where his/her hologram is, through very rich sensing of multiple sorts, synchronised to devices on his/her body, very close to the reality of body language and sensory experiences MR telepresence allows interaction with both physical and digital objects, near or far in physical reality.
- Non-material fashion - Using digital objects and overlays to create a personal expression that can be viewed and or otherwise sensed in MR by others. The user can choose who can see their digital outfits, and swap, sell and, purchase digital items from others and stores – as well as create his/her own outfits.
- Augmented shopping – In-store augmentation of products, expansion of assortment on display by MR, interactive presentations, and flat stores where everything is presented as touchable 3D objects. Will naturally also include sales of digital items and experiences.
- Interactive immersive advertisement – The digital space in the physical world will also be used as a new media, placing immersive interactive digital advertisements that tickle all your senses. The users should be able to control what type of advertisement, how much

and where they will be exposed to it, as well as if they agree to be part of an advertisement themselves. This also applies to business, brand, and real estate owners, as well as a government that should be able to block sensitive, dangerous, or important sites from advertisement.

- Multiplayer MR gaming – immersive gaming experiences together with others in public spaces could involve overlays of both environment and bystanders, player augmentation, digital objects, and non-player characters. The safety of the players and others in relation to traffic, people, and the public environment is a key issue, as well as the privacy and integrity of 3rd parties (including persons and objects).
- Guidance and information – wayfinding and city guides in MR, with a mix of different sensory experiences, such as sound, haptics, or fragrance. Any information about anything can be searched for by looking at, pointing to, or touching a physical thing.
- Social media sharing – Posting or streaming three-dimensional all senses experiences in social media, such as a concert, but also local social communication such as digital creations to be posted in cyber-physical space ("digital tagging"), which leads to similar requirements as for advertisements.
- MR micro-holidays – spending a few hours in the medieval version of your city centre, or by a telepresence experience in a quiet library in Oxford while you are on the train. MR can also be used for relaxing by shutting out annoying sounds and blocking out visuals.

2.3.3.1 Deployment characteristics

Table 2-6: ***Deployment characteristics for fully merged cyber-physical worlds***

Environment	Urban, primarily outdoor (street environments, parks, etc.) but also public indoor (malls, subways, etc.). High density of users.
Type of deployment	High density of cells with high bandwidth, with edge compute and spatial mapping, adapted for low latency
Users / devices	Number of devices per user can be 3-10 of which more than one may be connected directly to the network and others are locally connected to those devices. Connection density can be very high in crowded areas (train stations, malls), up to 3 persons m^2. Physically present and telepresence consumers with Augmented Reality (AR) glasses and body sensors/actuators (tactile gloves, electromyography (EMG) wristbands, smart watches, smart fabrics, etc.), and smartphone/connectivity puck.
Mobility	Highly nomadic but typically walking speed. Users can also be stationary mobile in buses and other city vehicles.
Frequency bands	Higher frequency bands in small cells with high bandwidth, mmWave, and higher
Environmental constraints	High cost of cell sites. Dense deployments cause interference issues
Any other constraints	Safety is a critical aspect when people move in city scenarios with traffic while interacting in digital worlds

2.3.3.2 KPI targets

Table 2-7: ***Target KPIs for fully merged cyber physical worlds***

		KPI	Target value	Reasoning / References
Communication	Dependability Attributes	Availability [%]	99	Acceptable with some service gaps, depends on the robustness of utilised codecs (i.e., impact on user-perceived quality should be low to avoid nausea).
		Reliability [%]	99.9	When QoS must be met to avoid nausea and user distress.
		Safety	Critical	Geo-fencing service must protect users from physical harm due to occlusion issues or shut service down.
		Integrity	High	Bystander integrity protection and protection against 3rd party misuse of usage data.
		Maintainability	Mid	Acceptable with some service downtime.
	QoS Attributes	Service latency [ms]	<20 ms	E2E roundtrip UL+DL less than 20 ms for at least 99% of the time. Should be <100ms for 99.99% of packets to avoid distress and discomfort if video frames are dropped due to late arrival (depending on codec).
		Data rate (minimum expected, desired, maximum) [Mbit/s]	1 Gbit/s DL 0.1 Gbit/s UL	Per user (from multiple devices): ~1 Gbit/s DL (AR stream 0.5 Gbit/s, spatial map 0.5 Gbit/s), ~100 Mbit/s UL (spatial map + user data).
		Resource efficiency		*Refer to deployment description (e.g., frequency, energy consumption)*
		Scalability		*Refer to deployment description (e.g., number of users, mobility, ...)*
AI and computation	Dependability Attributes	Agent availability [%]	99%	Same level as communication.
		Agent reliability [%]	99.9%	Same level as communication.
		Safety	Critical	Same level as communication.
		Integrity	High	Same level as communication.
		Maintainability	Mid	Same level as communication.
	QoS Attributes	AI service RTT [ms]	<20 ms 99%	Same level as communication.
		Inferencing accuracy [%]	N/A	No specific requirements.
		Interpretability level	N/A	No specific requirements.
		Training/model transfer latency [ms]	N/A	No specific requirements.
		Resource efficiency		*Refer to deployment description (e.g., frequency, energy consumption)*
		Scalability		*Refer to deployment description (e.g., number of users, mobility, ...)*

Localisation and Sensing	Dependability Attributes	Service availability [%]	99%	Sensing functionality must work for service to be safe.
		Service reliability [%]	99.99%	Considering the sensing functionality as a whole but not individual sensors.
		Safety	Critical	Same level as communication.
		Integrity	High	Same level as communication.
		Maintainability	Mid	Same level as communication.
	QoS Attributes	Location accuracy [m]	0.1	Application may need higher precision to place a digital overlay on surroundings but can improve the precision with image analysis. User's devices also need to be located relative to each other.
		L/S service RTT [ms]	50	Synchronisation of movement and a digital overlay.
		Orientation accuracy [°]	5	Application may need higher the precision to place digital overlay on surroundings but can improve precision with image analysis. User's devices also need to be oriented relative to each other.
		Refresh rate [1/s]	20	Synchronisation of movement and digital overlay.
		Minimum and maximum resolvable range [m]	0.1-10	No general upper bound (maximum range) is specified for this use case, depends on deployment.
		Angular resolution [°]	1	Giving ~15 cm resolution at 10 m distance.
		Velocity range [m/s]	0.1-10	Unclear if needed.
		Velocity resolution [m/s]	< 10	In the order of human mobility to accurately overlay rendered objects with the real world.
		Resource efficiency		*Refer to deployment description (e.g., frequency, energy consumption)*
		Scalability		*Refer to deployment description (e.g., number of users, mobility, ...)*

2.3.3.3 Quantification of key values, KVIs

The key value of fully merged cyber-physical worlds depends on what it is used for. In order to link the use case to the SDGs, different utilisations, applications, and functions need to be investigated separately. As with many use cases, there will be several utilisations which would be beneficial for humans and society, however there is also a risk of both unintended impacts as well as intended misuse (sometimes called misuse case or abuse case). Utilising the MR telepresence in consultations with medical doctors can improve the access to healthcare (SDG 3.8: "*access to quality health-care services*"), while enhancing teacher-student interactions can be linked to improved education quality and availability (SDG 4.1 & 4.3: "*access to education"*). Improving work collaborations can reduce the need to travel and will therefore be linked to reduced energy-related CO2 emissions (SDG 13: "*climate action*"). Socialising with family and friends as well as virtual traveling will also reduce the need to travel and increase the possibility to experience other countries more often without traveling there (SDG 8.9: "*sustainable tourism*") as well as other benefits such as increased quality of life. In parallel to reduced travel needs there is also a

possibility to dematerialise through the reduced building of e.g., new large arenas when some people choose to participate via MR when visiting sports or concerts. KVIs can be defined for the respective areas. For the healthcare example, it could be "percentage of the population for which access to state-of-the-art basic healthcare is providable over 6G".

2.3.4 Interacting & cooperative mobile robots & flexible manufacturing

The *interacting and cooperative mobile robots* use case as described in [HEX21-D12] contains aspects of *flexible manufacturing*. The focus is on the direct interaction between mobile robots on a factory shop floor, with robots acting towards a common goal (e.g., fulfilment of a production task). In addition to direct machine-to-machine interaction, humans can be involved or even required in some of these interactions. Interactions among robots and humans include:

- Multiple mobile robots carrying a good in-between modular flexible production cell, requiring accurate positioning and synchronisation among participating entities.
- Mobile robots interacting with a (fixed) machine as part of a modular flexible production cell (e.g., transporting goods to a machine, holding a production item).
- Humans interacting with machinery or mobile robots indirectly (e.g., by approaching them as sensed by the system) or directly (e.g., through mobile or mounted Human-Machine Interfaces (HMIs) or jointly working on the same production item).
- Robots learning from humans as a result of interaction on a task, e.g., optimizing execution steps or improving error mitigation and prevention steps.

We consider an indoor scenario in a production/manufacturing environment. The use case and its functional aspects is further described in [HEX21-D71].

2.3.4.1 Deployment characteristics

Table 2-8: ***Deployment characteristics of Interacting and Cooperating Mobile Robots***

Environment	Indoor use case (e.g., factory hall, shopfloor), consisting of a mixture of open space, warehouse zones, and static machinery grouped into production cells.
Type of deployment	Small-cell environment served by on-premises infrastructure.
Users / devices	Predominant machine-to-machine communication between robots, AGVs, and static machinery (density: ~1 per sqm), for collaborating robots up to 5 per sqm [HEX21-D71] AR/Virtual Reality (VR) or handheld devices for direct human/machine interaction on the shop floor. Indirect, intent-based interaction is caused by human behaviour (e.g., crossing a trajectory of an AGV, approaching a machine).
Mobility	Mobile robots / AGVs with movement speed <10 m/s [HEX21-D71]
Frequency bands	Focus on campus-wide (here: indoor shopfloor) reliable coverage, also in non-Line of Sight (LoS) conditions and with blockage. Potential for the utilisation of dedicated private frequency bands for industrial (campus) networks.

	Potential for direct device-to-device communication at higher frequency bands within local flexible production cells or among collaborating robots for ultra-low latency.
Environmental constraints	Challenging propagation environment, especially in warehouse areas or in the presence of large machinery. LoS blockage caused by material being moved across the factory floor.
Any other constraints	Sensitivity of production-related data and models and corresponding regulations or customer requirements. Predominantly brownfield deployments, i.e., existing machinery and, legacy (communication) infrastructure needs to be supported. Long lifetime of equipment, backward compatibility and upgradeability is expected. No (significant) constraints in terms of energy efficiency for collaborating robots. Energy efficiency is relevant for mobile robotics in general (e.g., AGVs), but communication budget is expected to play a minor role.

2.3.4.2 KPI targets

The target values stated in the table below correspond to the collaborative robots aspect of the envisioned Industry 4.0 use case, unless otherwise noted. Most target ranges are derived from [22.804] and [5GA19]. For related aspects in the wider scope of the industrial automation use case, please also refer to [22.104].

Table 2-9: ***Target KPIs for Interacting and Cooperative Mobile Robots***

		KPI	Target value	Reasoning / References
Communication	Dependability Attributes	Availability [%]	99.9999	Survival time is equal to the transfer interval of control packets [22.104], with periodic deterministic communication.
		Reliability [%]	up to 99.9999999	[HEX21-D71]
		Safety	critical	Regulations for product safety for machines apply. Signalling of alarms and incidents, geo-fencing for human-machine interaction.
		Integrity	critical	Protection against 3rd party usage or manipulation of (sensitive) production data
		Maintainability	high	Restoration of service is a key priority, also by non-5G/6G experts available on-site.
	QoS Attributes	Service latency [ms]	0.5 - 25	RTT (sensor to a controller to an actuator) of 1-50 ms for collaborating robots [HEX21-D71]. Deterministic communication is required for control applications.
		Data rate (minimum expected, desired, maximum) [Mbit/s]	< 0.1 for control	Per-UE control traffic is expected to be in the kbit/s range, with small packet sizes (<1.5 kByte)
		Resource efficiency		*refer to deployment characteristics (prev. section)*
		Scalability		*refer to deployment characteristics (prev. section)*

AI and computation	Dependability Attributes	Agent availability [%]	99.9999 for control, less otherwise	If used as part of the control algorithm, requirements as stated for the communication service apply. Requirements are less strict if AI is utilised for offline process optimisation or recommendations to human operators.
		Agent reliability [%]	up to 99.9999999 for control, less otherwise	C.f. reasoning for agent availability.
		Safety	critical for control	If utilised for controlling robots, safety constraints for human-machine co-working need to be guaranteed.
		Integrity	high	Protection against 3rd party usage or manipulation of production data (model) and training data.
		Maintainability	high	Required to recover within the required time to ensure application productivity. Depends on the area of utilisation of AI in the use case (e.g., online or offline optimisation of processes and the impact on productivity).
	QoS Attributes	AI service RTT [ms]	part of overall RTT budget of 1 – 50 ms	If utilised in control applications. Less strict for offline optimisation or recommendations to human operators.
		Inferencing accuracy [%]	high	Depends on utilisation of AI in the use case and respective quality function. Generally, accuracy is required to be high to lead to a realization of the expected benefits when utilising AI.
		Interpretability level	high	Especially in safety-related applications (e.g., in control applications), the explainability of utilised algorithms is a key requirement.
		Training/model transfer latency [ms]	-	Assuming offline training in the use case.
		Resource efficiency		*refer to deployment characteristics (prev. section)*
		Scalability		*refer to deployment characteristics (prev. section)*
Localisation and Sensing	Dependability Attributes	Service availability [%]	> 99.99	[HEX21-D31]
		Service reliability [%]		No use-case-specific requirements are known.
		Safety	critical	If used for safety-critical aspects (e.g., machine shutdown in case of human presence within close proximity). If used for monitoring and/or optimization purposes with less stringent requirements, safety might not be an issue due to limited/no impact.

		Integrity	high	Robustness against unintended or intended interference, especially if used in safety-critical scenarios.
		Maintainability		No use-case-specific requirements are known. Application productivity needs to be high.
	QoS Attributes	Location accuracy [m]	0.01 – 0.05	[HEX21-D31]
		L/S service RTT [ms]	0.1 – 100	Inverse of required update rate.
		Orientation accuracy [°]	sub-degree	[HEX21-D31]
		Refresh rate [1/s]	10 – 10000	Once per 100 ms to once per 0.1 ms
		Minimum and maximum resolvable range [m]		No use-case-specific requirements are known.
		Angular resolution [°]	sub-degree	[HEX21-D31]
		Velocity range [m/s]	0 – 10	Required accuracy is expected to vary for different velocities. Collaboration with high required location accuracy might only occur in situations with reduced velocity.
		Velocity resolution [m/s]	0.5	[HEX21-D31]
		Resource efficiency		*refer to deployment characteristics (prev. section)*
		Scalability		*refer to deployment characteristics (prev. section)*

2.3.4.3 Quantification of key values, KVIs

Linking the use case to the UN SDGs, there is a direct relation to SDG 9: "*industry innovation and infrastructure"*. With improved human-machine interaction and intelligent cooperation among robots, resource-use efficiency and overall sustainability will be increased (SDG 9.4). More complex tasks are solved by intelligent cooperation and interaction of existing machinery, benefitting from human knowledge and problem-solving skills, also contributing towards SDG 8 (SDG 8.2). The manufacturing value is expected to increase as a consequence of inclusive (cooperative and interactive) industrialisation (SDG 9.2). The ability to have a more flexible and adaptive manufacturing process contributes to this increase in manufacturing value and, potentially, the overall energy efficiency (SDG 7.3). Utilising interactive and cooperative mobile robots for complex and highly customised tasks lowers the entry barrier for small-scale industrial enterprises (SDG 9.3). Ensuring safety in human-machine interaction by utilising novel 6G capabilities such as localisation and sensing contributes towards safe and secure working environments (SDG 8.8).

The increase in resource-use efficiency in industrial processes gained from mobile and flexible production can serve as a direct indicator for the "6G for sustainability" Hexa-X key value. Impact on the overall supply chain and potential reduction of waste due to more customisable production remains challenging to be quantified, but also contribute to this key value and the SDG 12.5.

2.3.5 Immersive smart cities & integrated micro-networks for smart cities

Immersive smart cities, as well as Integrated micro-networks for smart cities, are two use cases where a very large number of IoT devices and sensors are monitoring the different flows in the city for succeeding the best operation in all the city sectors. The former use case belongs to the massive twinning use case family, relying on information gathered from various sensors to provide a 4D map of all different city flows. The latter use case belongs to the hyperconnected resilient infrastructures use case family and builds upon the possibility to leverage multiple types of IoT networks, with different ownerships and nature, to collect information and share part of the infrastructure.

2.3.5.1 Deployment characteristics

Table 2-10: ***Deployment characteristics for immersive smart cities and integrated micro-networks for smart cities***

Environment	Urban and probably suburban mainly outdoor environment for transportation, environment and safety sectors of the city and mainly indoor environment for education and healthcare sectors of a city.
Type of deployment	The type of deployment varies with the use cases, as different use cases will involve different types of deployment areas. For instance, macrocells might be more appropriate for transportation use case and micro cells for the personal healthcare use case.
Users / devices	In most cases, there needs to be an interaction between different segments and components such as end-user devices, small cells, edge cloud, core, and cloud. The number of embedded devices might be up to trillions and connection density might be up to 10^7 devices/km^2 depending on the city [HEX21-D71].
Mobility	Users and devices in these use cases are mobile in the case of transportation, roads, etc., and static in the case of the environment, energy, water, gas, etc.
Frequency bands	mmWave bands for extra capacity in dense areas. Sub-6 GHz for better coverage.
Environmental constraints	Challenging propagation environments such as underground roads, highways, and infrastructure in general since it can be difficult in those environments to obtain the right data flow and coverage for updating the digital twins among others. Dense deployments in some areas of large cities can cause interference issues.
Any other constraints	Trustworthiness of the network is crucial because of the vast amount of personal and public data collected and analysed for the needs of these use cases. The three indicators for KVIs of trustworthiness, safety, confidentiality, and integrity, should be measured and improved during the 6G research period. Dependability on the other hand, which is related to the coverage of the network, the latency of data transmissions, and probability of error, is a key quality for managing city flows.

2.3.5.2 KPI targets

Table 2-11: ***Target KPIs for immersive smart cities and integrated micro-networks for smart cities***

		KPI	Target value	Reasoning / References
Communication	Dependability Attributes	Availability [%]	99 - 99.99	Higher target value for use-cases such as localisation of control utilities and traffic monitoring. Lower target value for use-cases such as landscape sensing and digital twins of smart buildings. [HEX21-D31]
		Reliability [%]	99.999	Related to digital twin use case taken from [5GP21]
		Safety	high	Signalling of incidents in infrastructure city sector (roads, railways, buildings, transport, energy, water, etc.), environment city sector and healthcare city sector as well as safety/stability city sector where this KPI is of great interest.
		Integrity	critical	Protection against 3rd party usage or manipulation of sensitive production data.
		Maintainability	medium - high	Restoration is crucial in most of the city sectors' use cases.
	QoS Attributes	Service latency [ms]	0.1 - 100	Depending on the use case. Lower target value for indoor use cases and higher target values for outdoor use cases. Tolerable Jitter is expected to be 10% of the targeted service latency depending on the use case.
		Data rate (minimum expected, desired, maximum) [Mbit/s]	10 - 100	[HEX21-D71]
		Resource constraints		*refer to deployment characteristics (prev. section)*
		Scalability		*refer to deployment characteristics (prev. section)*
AI and computation	Dependability Attributes	Agent availability [%]	99 – 99.99	Same as communication
		Agent reliability [%]	99.999	Same as communication
		Safety	critical	Especially when considering human-machine interactions for instance in the healthcare city sector.
		Integrity	critical	Protection against 3rd party usage of AI models and (sensitive) training data.
		Maintainability	medium - high	Depends on the area of utilisation of AI.
	QoS Attributes	AI service RTT [ms]		Depends on the application.
		Inferencing accuracy [%]		Depends on the application.
		Interpretability level	high	Especially in safety-related applications.

		Training/model transfer latency [ms]	N/A	Assuming offline training mostly.
		Resource constraints		*refer to deployment characteristics (prev. section)*
		Scalability		*refer to deployment characteristics (prev. section)*
Localisation and Sensing	Dependability Attributes	Service availability [%]	Up to 99.9999	[HEX21-D71]
		Service reliability [%]	99.999	Same as communication
		Safety	critical	Same as communication.
		Integrity	critical	Robustness against unintended or intended interference.
		Maintainability	medium - high	Same as communication.
	QoS Attributes	Location accuracy [m]	< 1 m horizontal and vertical for utility services such as transportation, piping, garbage etc. < 0.025 m for future smart buildings (half thickness of typical wall)	[HEX21-D31]
		L/S service RTT [ms]		No use-case-specific requirements are known.
		Orientation accuracy [°]		No use-case-specific requirements are known.
		Refresh rate [1/s]	1/3600 - 1	Once per hour for use cases such as digital twins of smart buildings, once per minute for use cases such as landscape sensing and, once per second for use cases such as traffic monitoring and localisation of control utilities. [HEX21-D31]
		Minimum and maximum resolvable range [m]		No use-case specific-requirements are known.
		Angular resolution [°]		No use-case-specific requirements are known.
		Velocity range [m/s]	-20 m/s to 20 m/s	For use cases such as traffic monitoring. [HEX21-D31]
		Velocity resolution [m/s]	0.5	For use cases such as traffic monitoring. [HEX21-D31]

		Resource constraints		*refer to deployment characteristics (prev. section)*
		Scalability		*refer to deployment characteristics (prev. section)*

2.3.5.3 Quantification of key values, KVIs

These two use-cases, Immersive smart cities and Integrated micro-networks for smart cities, are closely linked to UN SDG 11 "*Make cities and human settlements inclusive, safe, resilient and sustainable*". However, on a more detailed level, there are several other SDGs and SDG targets that could be linked, depending on what the solutions are used for. As one example, a 4D map of a water and sewage system, could be directly linked to SDG 6.4: increase water use efficiency, SDG 6.5: Implement integrated water resources management and SDG 9.1: Develop quality, reliable, sustainable and, resilient infrastructure, and, by extension, also to health-related SDG targets associated to water-borne diseases. Another example is a solution used to monitor and control traffic flows which could both be related to safety (SDG 3.6: halve road traffic accidents) and decrease the greenhouse gas emissions (SDG 13: climate action) by optimizing the traffic flow. For smart grid and microgrid scenarios, the use case targets SDG 7.3, improvement of energy efficiency.

2.3.6 Infrastructure-less network extensions and embedded networks

Infrastructure-less network extensions enable connectivity where the infrastructure-based connectivity is temporary or locally not available, not sufficiently performing, or inefficient from an energy consumption point of view while all involved nodes would in principle be able to connect to the infrastructure. Examples are platoons of agriculture vehicles being interconnected in a harvesting campaign or groups of construction vehicles in a road construction scenario. In both cases, in particular in rural areas coverage might be limited while the interconnection should be preserved when the coverage area is left. Embedded networks are networks of sensors in machines, vehicles, or other environments where the environment provides a common functional context, e.g., all sensors belonging to a train and supporting specific safety or operational functionalities. Not all but some of these sensor nodes are able to connect to the infrastructure. Sensor nodes are often deployed by one Operational Technology (OT) manufacturer and belong to machines or vehicles of the same manufacturer.

2.3.6.1 Deployment characteristics

Table 2-12: ***Deployment characteristics for infrastructure-less network extensions and embedded networks***

Environment	Use cases can be found almost everywhere. However, typical scenarios are in rural areas where coverage is limited or not always available with the required levels of performance (e.g., data rate, latency...). In production scenarios, larger populations of machines, modules, or vehicles of the same vendor might be interconnected by infrastructure-less networks forming an underlay network to the infrastructure network. Embedded networks can be found in vehicles, trains, AGVs, machines, planes, ships, and as body area networks.
Type of deployment	Sensor networks with sensor-to-sensor connectivity.

Users / devices	Sensor populations in machines, vehicles, trains, robots, ships, or planes are typical examples for embedded and partly infrastructure-less networks of some tens to some thousands of sensors. Another example is Body Area Networks (BAN) consisting of biosensors for medical and well-being applications that can also be seen as examples for infrastructure-less networking. In other cases, machines, vehicles, etc., are interconnected mutually The size of these populations ranges from several sensors to some hundred or thousand sensors. The sensor density is use case-specific and differs between some sensors per cubic meter (BANs, embedded network in machines) to even lower values (interconnected vehicles in construction or agriculture campaigns).
Mobility	The majority of use cases consist of either a group of sensor nodes in the same environment and thus moving on highly correlated trajectories (but with potentially high-speed relative to the infrastructure nodes). Here, the nodes have little or no mutual distance variation. Or, in other cases, the population of nodes forming the infrastructure-less network are moving independently but in close proximity allowing direct sensor-to-sensor connectivity.
Frequency bands	The sensor population are expected to share a standard licensed spectrum below 6 GHz with the infrastructure. Higher frequency bands might be used as well, but bands below 6 GHz will probably be preferred due to the reduced power consumption in low bands.
Environmental constraints	The implementation in machines and in body area networks will limit the form factor of the sensors. In some cases, as body area networks or embedded networks in machines, a close proximity of sensors is possible. EMC requirements will result from the use in medical and machinery scenarios. In production environments, specific safety of operation requirements has to be taken into account as well.
Any other constraints	Cost and form factor of sensors, flexibility of spectrum usage, Variable authentication methods

2.3.6.2 KPI targets

Table 2-13: ***Target KPIs for infrastructure-less network extensions and embedded networks***

		KPI	Target value	Reasoning / References
Communication	Dependability Attributes	Availability [%]	critical	Often part of safety-related applications, see [HEX21-D71]
		Reliability [%]	Very high	Often part of safety-related applications
		Safety	critical	Often part of safety-related applications
		Integrity	Very high	Often part of safety-related applications and accessible by third parties
		Maintainability	Very high	Often part of professional applications

	QoS Attributes	Service latency [ms]	1 ms	Assuming that Time-Sensitive Networking (TSN) applications are not covered. Can be significantly higher for some use cases where multi-hop communication and/or energy preserving strategies are utilised.
		Data rate (minimum expected, desired, maximum) [Mbit/s]	Kbit/s up to some 10 Mbit/s	Video
		Resource constraints		*refer to deployment characteristics (prev. section)*
		Scalability		*refer to deployment characteristics (prev. section)*
AI and computation	Dependability Attributes	Agent availability [%]	Very high	Safety-critical applications
		Agent reliability [%]	Very high	Safety-critical applications
		Safety	N/A	
		Integrity	High	
		Maintainability	N/A	
	QoS Attributes	AI service RTT [ms]		Depending on the application, no use case-specific requirements are known.
		Inferencing accuracy [%]		Depending on the application, no use case-specific requirements are known.
		Interpretability level		Depending on the application, no use case-specific requirements are known.
		Training/model transfer latency [ms]		Depending on the application, no use case-specific requirements are known.
		Resource constraints		*refer to deployment characteristics (prev. section)*
		Scalability		*refer to deployment characteristics (prev. section)*
Localisation and Sensing	Dependability Attributes	Service availability [%]	Very high	Manoeuvring and safety applications
		Service reliability [%]	Very high	Manoeuvring and safety applications
		Safety	Very high	Manoeuvring and safety applications
		Integrity	Very high	Manoeuvring and safety applications
		Maintainability		Depending on the application, no use case specific requirements known.
	QoS Attributes	Location accuracy [m]	0.01	Shopfloor, agriculture
		L/S service RTT [ms]	1 - 10	

		Orientation accuracy [°]	1, 6D	
		Update rate [1/s]	100 - 1000	
		Minimum and maximum resolvable range [m]	0.01 - 1	
		Angular resolution [°]	1	
		Velocity range [m/s]	0.1 - 100	
		Velocity resolution [m/s]		Depending on the application, no use case-specific requirements are known.
		Resource constraints		*refer to deployment characteristics (prev. section)*
		Scalability		*refer to deployment characteristics (prev. section)*

2.3.6.3 Quantification of key values, KVIs

Use cases can be linked to the following UN SDGs:

- SDG 9: "*industry innovation and infrastructure*" – improved connectivity between machines, vehicles, and modules enables digitalisation, new business models, and a more efficient operation.
- Efficiency of resource usage and, with this, overall sustainability will be improved (SDG 9.4).
- SDG 8.8 is addressed by enabling robust safety mechanisms that are independent of infrastructure when needed.
- Indirectly SDG 12.5 is supported as the overall digitalisation in supported in many domains and, with this, an increased resource efficiency, an overall circular economy, and a reduction of waste can be achieved.

3 E2E architecture

The main objective of this chapter is to summarise the technical enablers which are required for the 6G architecture envisioned by Hexa-X. The technical enablers are the important components for the transformation to the new architecture and they are essential for supporting the requirements of 6G use cases presented in the previous chapter as well as in D1.2 [HEX21-D12]. To this end, in this chapter, the most impacting architectural enablers are thoroughly reviewed and their requirements from the new architecture are also characterised.

This chapter is organised as follows: Section 3.1.1 describes the architectural principles. An overview of the proposed end-to-end (E2E) 6G architecture is demonstrated in Section 3.1.2 Thereafter the requirements and impact of RAN technologies and localisation and sensing are defined in Section 3.2. Subsequently, the necessary 6G enablers for intelligent, flexible, and efficient networks are discussed in Sections 3.3, 3.4, and 3.5, respectively. Finally, Section 3.6 covers enablers for Service Management and Orchestration.

3.1 Introduction to E2E architecture

The following section briefly describes the architectural principles. A preliminary E2E 6G system view architecture is proposed in Subsection 3.1.2 as well as a brief description of individual enablers and layers. This section also touches on the security enablers required for the 6G architecture. A more detailed description can be found in Chapter 4, dedicated to security, privacy. and trustworthiness topics.

3.1.1 Architectural principles

In [HEX21-D51] eight different architectural principles were defined. These principles should be fulfilled by the design of 6G architecture. A summary of architectural principles can be found in Figure 3-1.

Principle 01: Exposure of capabilities

The architecture solution shall expose new and existing network capabilities to E2E applications and management such as predictive orchestration.

Principle 02: Designed for (closed loop) automation

The architecture should support full automation to manage and optimise the network without human interaction.

Principle 03: Flexibility to different topologies

The ability of the network to adapt to various scenarios without loss of performance while still enabling easy deployment.

Principle 04: Scalability

The network architecture needs to be scalable both in terms of supporting very small to very large-scale deployments, by scaling up and down network resources based on needs.

Principle 05: Resilience and availability

The architecture shall be resilient in terms of service and infrastructure provisioning using methods such as multi-connectivity and separation of Control Plane (CP) and User Plane (UP), removing single points of failure.

Principle 06: Exposed interfaces are service based

Network interfaces should be designed to be cloud-native, utilising state-of-the-art cloud platforms and IT tools in a coherent and consistent manner.

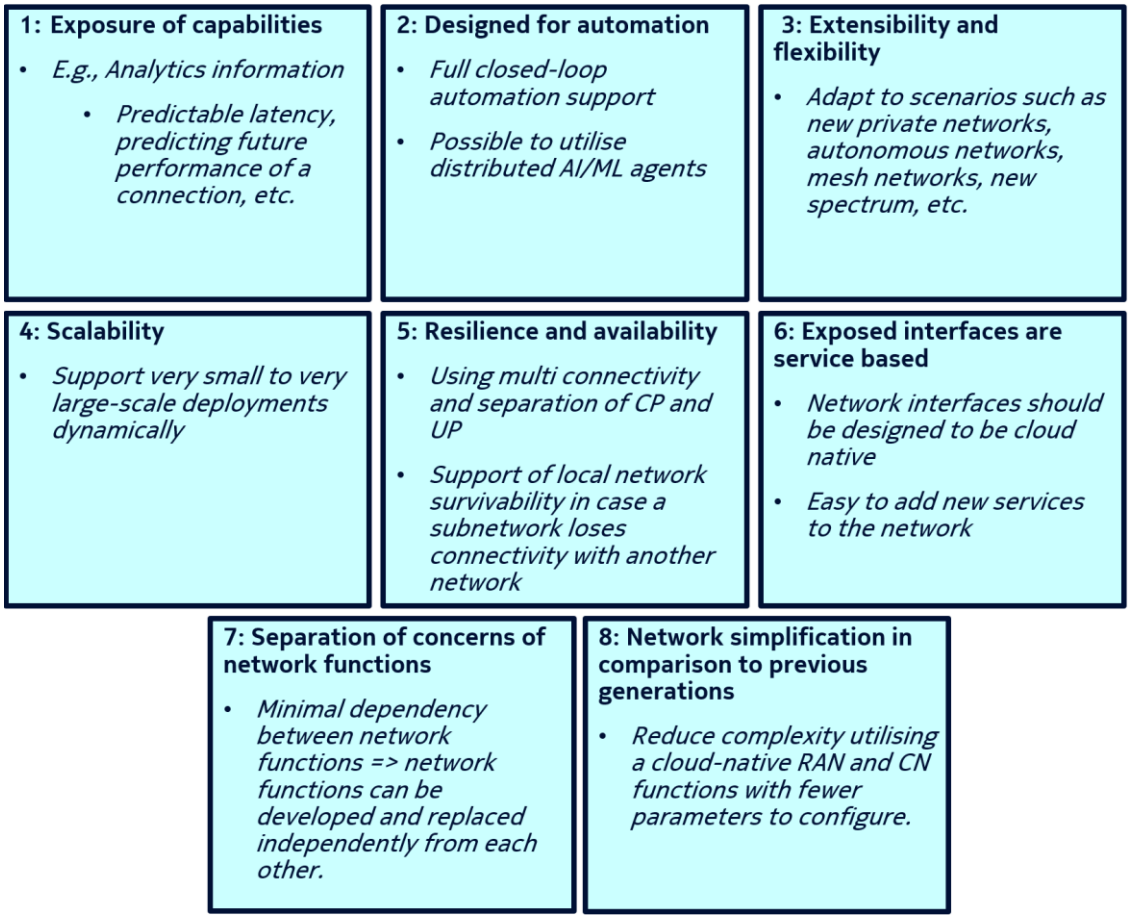

Figure 3-1: ***6G architecture principles from [HEX21-D51], guiding the architecture design***

Principle 07: Separation of concerns of network functions

The network functions have bounded context and all dependencies among services are through their APIs with a minimal dependency with other network functions, so that network functions can be developed, deployed and replaced independently from each other.

Principle 08: Network simplification in comparison to previous generations

Streamline the network architecture to reduce complexity utilising cloud-native upper layer RAN and CN functions with fewer (well-motivated) parameters to configure and fewer external interfaces.

A more detailed description of the architectural enablers and how they intend to fulfil the 6G architectural principles can be found in the following Sections 3.2 to 3.6 and respective technical Hexa-X deliverables.

3.1.2 Hexa-X E2E architecture overview

Figure 3-2 depicts a high-level view of the Hexa-X E2E 6G architecture and highlights the key technical enablers. The various building blocks are organized into three layers: Infrastructure, Network Service, and Application.

The infrastructure layer is comprised of the Radio Access (RAN), Core (CN), and transport Networks which contain radio equipment (non-virtualised radio functions like RUs, DUs, or even classical base stations), switches, routers, communication links, data centres, cloud infrastructure, and so on. The infrastructure layer provides the physical resources to host the network service

and application layer elements. Furthermore, due to the introduction of new use cases, e.g., immersive smart city [HEX21-D12], the infrastructure layer envisioned for 6G can accommodate new enablers such as localisation and sensing, see Section 3.2.3. A thorough gap analysis was conducted on localisation and sensing in deliverable Hexa-X D3.1 [HEX21-D31]. The infrastructure layer also contains RAN improvements for extreme low latency, high reliability, and availability. More details on the evolution of RAN technologies can be found in Sections 3.2.1 and 3.2.2 as well as in the deliverables D2.1 [HEX21-D21] and D2.2 [HEX21-D22]. The 6G architecture incorporate different (sub)network solutions into a network of networks. The network of networks can easily and flexibly adapt to new topologies to meet the requirements of both extreme performance and global service coverage, well beyond what 5G is capable of, see Section 3.4.

The network service layer is envisioned to be entirely cloud-based with function and microservices expanded from central cloud to extreme edge cloud. By having all network functions, operations, and applications implemented as microservices, we can move toward a softwarised, intelligent and efficient 6G architecture. Section 3.3 on enablers for intelligent network describes mechanisms to support AI in 6G and AIaaS, programmability, and network automation. Further on, with a cloud-native approach, the RAN and CN architectures can be streamlined, e.g., reduce some complexity by removing multiple processing points for a certain message and removing duplication of functionalities among functions. This topic is further investigated in the section related to enablers for efficient networks. (see Section 3.5).

One of the key technology enablers of the network service layer is the introduction of the extreme/far edge cloud. Extreme edge cloud covers the part of the network with high heterogeneity of devices with a wide variety of technologies, in terms of both hardware and software. These devices could be personal devices (smartphones, laptops…), and a huge variety of IoT devices (wearables, sensor networks, connected cars, industrial devices, connected home appliances, etc.). The concepts of edge and far-edge computing become more and more relevant for the 6G architecture and services. Cloud-native technologies will be required to create cloudlets at the edge of the network, with application-to-application and function-to-function communication capable to satisfy a large number of interconnected assets with flexible mesh topologies, see Section 3.6.

Another important aspect of this layer is the exposure framework and integration fabric. They establish a communication channel between multiple domains, enabling seamless interoperation and networking across different domains. More details can be found in Section 3.3.4.

The network management and orchestration are gradually moving toward increasing the levels of automation and fully automated closed-loop control. This is supported by the parallel adoption of advancements in Artificial Intelligence (AI) and Machine Learning (ML) technologies.

The aim is to provide a framework to optimally support reliability, flexibility, resilience and, availability through the concept of "orchestration continuum" - i.e., seamless orchestration spanning device-edge-cloud addressing changes in the infrastructure, requirements and failures.

Security and privacy mechanisms are an integral part of the overall architecture, affecting all network layers as well as the management and orchestration domain. Figure 3-2 highlights the 6G security technology enablers identified in [HEX21-D12] and further discussed in Chapter 4.

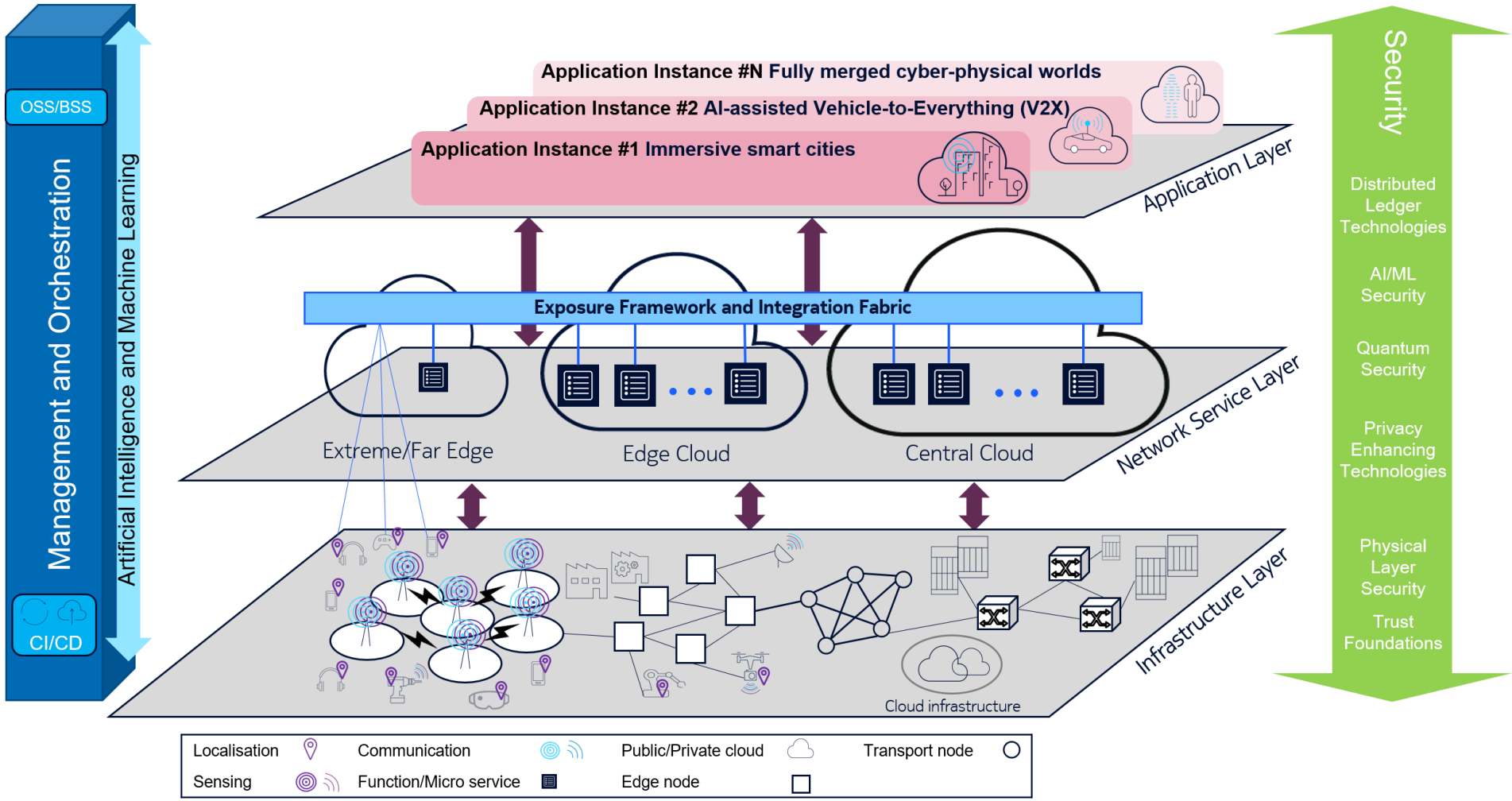

Figure 3-2: ***6G E2E architecture overview***

Trust foundations, physical layer security, and quantum security (in the sense of making use of physical quantum properties to support security, with quantum key distribution as the most prominent example) relate to the infrastructure layer. Quantum security in the sense of quantum-safe cryptography (i.e., cryptography being safe against quantum computer attacks) relates to cryptographic procedures on all layers including the management domain, for secure communication as well as secure storage. Privacy-enhancing technologies are important on all layers where sensitive data are gathered or processed, and clearly also in the management domain. Similarly, AI/ML security is relevant for all functions making use of AI/ML, in the sense of specifically protecting this use, but also refers to AI/ML-driven security mechanisms, e.g., in the management domain. Finally, distributed ledger technologies are relevant wherever it is required to establish "distributed trust", i.e., trust that is not anchored in a central trusted authority, as it may be the case in interdomain management, to give an example.

3.2 Enablers for RAN technologies and localisation

The main target of this subsection is to introduce three of the main RAN technologies which are necessary for the set of use cases introduced in the previous chapter. Technologies such as extreme high data rate radio links and Distributed large MIMO were investigated in detail in D2.1 [HEX21-D21] and D2.2 [HEX21-D22].

Localisation and sensing play an important role for most Hexa-X use cases to let these features become an integral part of next generation mobile communication. The following subsections summarise requirements that must be considered and integrated into the E2E architecture design.

3.2.1 Extreme high data rate radio links

Extreme high data rate links will be required in some very high-performance applications anticipated in 6G. Mostly those are related to highly advanced on-line imaging including holographic communications as well as providing extreme data rates for high-capacity cells. In those cases, a throughput of 100 Gbit/s or even significantly higher can be required. This means bandwidths of several tens of GHz as anticipated in reports D2.1 [HEX21-D21] and D2.2 [HEX21-D22] of this project. The architecture design, in particular the infrastructure layer, needs to ensure that such data rates can be brought to local small-scale base stations that will serve end-users. From network architecture point of view, this is not only implementing optical, or wireless backhaul connectivity with low latency, but it means a backplane that can support data rates of hundreds of Gbit/s on large scale. This is an expanded requirement for new use cases in 6G.

3.2.2 Distributed large MIMO

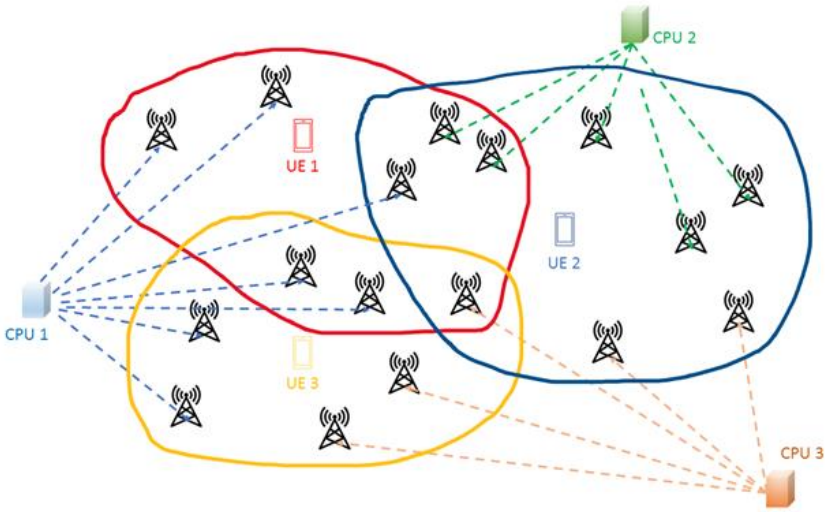

Figure 3-3: ***Illustration of distributed MIMO***

Distributed large MIMO (D-MIMO) (D2.1 Sec. 3.6 [HEX21-D21], D2.2 ch. 6 [HEX21-D22]) is a promising technology to address challenges in dense deployments at both low (cmW, lower mmW) and high (upper mmW and (sub-)THz) carrier frequencies, cf. (D2.1, Table 1-1 [HEX21-D21]). D-MIMO has the potential to

- allow for further densification of Access Points (APs) for increased and consistent area capacity.
- mitigate unreliable links due to shadowing/blockage thanks to macro diversity.
- achieve sufficient link margin despite output power limitations and high pathloss at upper mmW and (sub-)THz frequencies.
- allow for lowering Effective Isotropic Radiated Power (EIRP), simplifying deployment.

As illustrated in Figure 3-3, in D-MIMO UEs can be served by several APs that are controlled by one or several Central Processing Units (CPUs) via fibre-optic or wireless backhaul/fronthaul links. Wireless backhaul/fronthaul links can be implemented using dedicated frequency bands or using the same bands as for access, so-called Integrated Access Backhaul (IAB). As such, D-MIMO systems can implement various levels of cooperative MIMO systems ranging from Distributed Antenna Systems (DAS) to Joint Transmission Coordinated Multi-Point (JT-CoMP), (D2.1 Sec. 3.6 [HEX21-D21], D2.2 Sec. 6 [HEX21-D22]). When APs can perform channel estimation and distributed precoding locally, D-MIMO constitutes a scalable way to implement the network MIMO concept using distributed massive MIMO, also denoted as cell-free massive MIMO, [IFL19].

There is basic support available for the implementation of D-MIMO in 3GPP 5G standards (e.g., related to multi-TRP support). However, there are major gaps between theory and practical solutions on real-world deployment of D-MIMO, related to architecture and functional split between CPU(s) and APs, fronthaul/backhaul solutions, scalability, and efficient precoding techniques. Key conclusions from the studies so far within Hexa-X WP2 is that there are substantially different challenges and opportunities for D-MIMO at lower and upper frequency bands, calling for a scalable approach based on digital and analog solutions. That work also emphasises the need for efficient backhaul/fronthaul by integrating fibre and in-band wireless solutions. Since densification is the key enabler to meet coverage and reliability targets at the higher frequency bands and as it seems there is sufficient spectrum available, low-cost solutions are more important than spectral efficiency (at least in the early roll-out phases), cf. (D2.2 ch. 6 [HEX21-D22]). This calls for decentralised solutions at the higher frequency bands. In the lower frequency bands, the need for higher spectral efficiency calls for less distributed more digital approaches for better resource utilisation.

3.2.2.1 Impact on E2E architecture

Related to the 6G architectural principles described in Section 3.1.1, D-MIMO puts requirements in particular on

Extensibility and flexibility

D-MIMO systems can contribute to 6G systems being able to adapt to various scenarios. In particular, D-MIMO can implement a service to actively shape the propagation environment (cf. passive shaping using Reconfigurable Intelligent Surfaces (RIS)) for rank and multipath control towards programmable propagation environments, which might be very beneficial in certain scenarios. D-MIMO systems can also implement support for services provided by other network functions such as channel sounding for localisation, RF environment mapping, and multi-static sensing (radar) services. Due to the dense deployment, energy efficiency in D-MIMO APs is important. To this end, dynamic activation of AP functionality with short delay should be supported. The delay requirement would be driven by the need of the use case versus the activation/deactivation efficiency gains in the APs.

On the other hand, as input to enable efficient beamforming, shadowing/blocking mitigation and, resource allocation in D-MIMO systems, various context and situational awareness information would be beneficial, such as location, mapping and, dynamic sensing information.

Scalability

One of the key challenges for D-MIMO systems is scalability. Important network architectural enablers would be frequency agility, i.e., support for high (upper mmW) as well as low (cm and mmW) carrier frequencies, and flexible roll-out aspects as described in the following.

In the network architecture, there should be support for heterogeneous nodes (APs, compute & sensing nodes) having specialised functional roles such as supporting low latency communications, uplink RF processing, downlink RF processing, baseband processing, sensing, etc. There should be support for APs having heterogeneous HW capabilities related to transmitting power, carrier frequencies, processing, etc., and also various functionality, e.g., within the control plane (system broadcast, initial access, etc.), and within the user plane (unicast, multicast, targeted flow KPIs, etc.). Furthermore, the network architecture should support a various degree of CPU-AP functional split, and AP cluster sizes.

Dynamic scalability would also be important, such as flexible dynamic use of AP resources, various degree of AP network control capabilities (broadcast, paging, random access) and, UE idle modes (cell search, idle mode mobility, …). Various degrees of silence/sleep modes would be needed to maximise energy efficiency under service constraints. Support for pro-active resource allocation for highly mobile users and physical network slicing would further support dynamic scalability.

Resilience and availability

Multi connectivity and separation of CP and UP goes hand in hand with scalable D-MIMO systems. In particular, there should be support for joint multiband transmission/reception. In addition, management and orchestration functionality should operate on a sub-second time scale.

Separation of concerns of network functions

D-MIMO systems can offer capabilities for both communications and localisation and sensing services, and also benefit from localisation and sensing for optimising communications. Thus, a network architecture that supports separations of concerns with clear APIs is important. That would enable scalability, adaptability, support for heterogeneous communications and fusion of localisation and sensing information, and potentially also reuse of various training data (such as reference symbols).

Network simplification in comparison to previous generations

D-MIMO systems would benefit from an architecture that natively supports cloud-RAN.

3.2.3 Requirements for localisation and sensing

Localisation of User Equipment (UE) in mobile communication has been supported from the early stages of 3GPP. With 5G and its target use-cases, localisation is increasingly gaining importance [22.261]. Hexa-X and its visionary scenarios continue this trend and look at localisation that is even more accurate and has even stricter latency requirements (for more information regarding requirements and use cases see deliverable D3.1 [HEX21-D31]). Besides localisation, the technologies explored in Hexa-X open up the possibility of using next generation mobile communication system itself for sensing. Sensing use cases address for example the detection of landmarks by the network as well as locating humans even though not carrying any device (e.g., UEs). Localisation and sensing can become an inherent feature in next-generation mobile communication, but to meet the challenging performance goals, it must be an integral part of the system architecture.

Highly accurate localisation of UEs (Figure 3-4 A), as well as highly accurate localisation of assets (Figure 3-4 B), are examples of Hexa-X scenarios. Totally new scenarios that are expected are radar-like sensing scenarios where next-generation mobile communication devices can detect and track objects or humans that do not carry any device (Figure 3-4: C), and even gesture detection of humans without UEs is conceivable. The double-coloured radio waves in Figure 3-4 depict the combination of communication and sensing (overlap between blue and purple area): *joint radar, communication, computation, localisation, and sensing (JRC2LS).* Certainly, pure communication or pure sensing/localisation scenarios must be made possible in the future as well, depending on the application requirements.

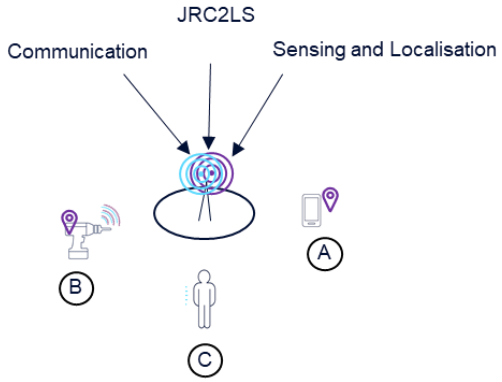

Figure 3-4: ***Examples of localisation and sensing scenarios (A: UE localisation, B: asset localisation and C: Detection of assets and humans without UEs via sensing)***

Flexible switching and prioritisation between pure communication, pure sensing/localisation, and a combined JRC2LS service capability should be considered in next generation mobile communication E2E architecture. Note that all three cases are envisioned to share the same hardware.

In the future, localisation and sensing should be designed as base functions or microservices. Accessing information from localisation and sensing services should be possible at different processing stages (e.g., raw sensing data as well as readily calculated position information) via the **exposure framework**. Interfaces for localisation services will need to be extended, e.g., from 3D to 6D (3D position + additional 3D orientation) and, totally new services, protocols and interfaces must be developed for sensing features.

Depending on the (access) rights of service consumers, access to information shall be possible or prohibited. Position and sensing data often are very sensitive data as they can easily be linked to personal information or business/trade secrets which must be protected from misuse. In industrial scenarios, this might even mean that such kind of data must never leave the factory. Localisation and sensing information will be also generated by the mobile communication network**. Securing this information** will be a very important architectural design requirement as the mobile network is not only passing information from application to application through the network but generating sensing information itself. This generated localisation and sensing information must be correct and trustworthy.

Low latency, which is understood as a short duration between the initialisation of sensing/localisation procedure and acquiring a localisation/sensing estimate, is also a challenge for the E2E architecture. But in general, service consumers must be able to describe applications' functional and non-functional localisation and sensing requirements like latency or reliability towards next generation mobile communication services and must be able to rely on these agreed quality parameters. These parameters are not necessarily static and might change over time which requires **flexible Quality of Service contracts.**

3.3 Enablers for intelligent network

The purpose of the Intelligent Network of Hexa-X architecture is to define the underlying mechanisms to support embedded AI for 6G, the concept of AIaaS introduced in [HEX21-D12], and to ensure dynamic adaptability of the network architecture to new use cases while keeping the infrastructure and energy costs at acceptable and sustainable levels. Built-in and integrated AI/ML depend on intelligence and automated closed-loop network operations to satisfy any E2E service KPIs.

The ultimate target for Intelligent Network is to enable autonomous networks, with no (or minimal) human intervention leveraging cognitive, closed-loop control network functions that can be instantiated on an on-demand basis even across network domain boundaries. Therefore, the constituent network functions need to be self-adaptable to new environments for which they were not originally planned for. Intelligent Network integrate different software implementations of AI functionality, multiple AI-agent setups and, different learning architectures with AI-driven network orchestration and cross-domain function placement, and built-in data analytics frameworks. To ensure network function adaptability we need also to consider programmability as part of the toolset of architectural adaptability. With this toolset new types of functionality could be added, and unnecessary features removed to keep the system concise and sustainable. Network programmability extends the network automation to the lowest level of functional granularity by offering mechanisms and interfaces to fast functionality exploration, updates, and component reconfiguration. This calls for corresponding capabilities on the UE side through UE programmability.

In the following Sections (3.3.1 – 3.3.4) we cover the requirements for each of the identified enablers for adaptable, AI/ML-integrated autonomous network architecture starting from the role of network and UE programmability, followed by AI-driven network automation that lays out the supporting machinery for the AIaaS framework. We consider AIaaS and its characteristics and assess the needs for new AI-enabled protocols which is then followed by network function placement that provides operational flexibility across multiple domains.

3.3.1 UE and network programmability

In Hexa-X we consider programmability in a wider architectural context in comparison to the well-known Software-Defined Networking (SDN) approach focusing on network and transport protocols or the application "plug-ins" extending application features. Programmability at the architecture level is seen as the basic means to adapt the architecture to new operational environments and to optimise network performance to diverse use cases and their KVIs/KPIs. It allows a faster pace of innovation and adaptability to new environments on the device and network node levels. The early approach to network programmability through SDN led to the separation of control and user planes and the introduction of specific SDN controllers at the network control plane. SDN approach is suitable to make the control of transport networks adaptable and easy to add new features. However, combining SDN with the softwarisation of network functions and cloudification trend increase further the potential of programmability. Softwarisation and cloudification have led to the introduction of new abstractions of infrastructure and network functions suitable for software developers and alignment with Continuous Integration/Continuous Delivery (CI/CD) pipeline, such as Service-Based Architecture (SBA) of the 5G core network with cloud-native APIs. Programmability combined with the AI/ML-capabilities and closed-loop control is a powerful tool to make the full E2E architecture adaptable to various environments, but it sets also new requirements on how the architecture should be constructed, how

Network Functions (NFs) would be designed and orchestrated and what kind abstractions and corresponding APIs need to be exposed by the physical infrastructure (e.g., generic APIs for accelerators) and the management system (e.g., version control, security, and compliance to standards). UE and network programmability mirror each other and provide a new tool for network automation and AI/ML functions.

3.3.1.1 Programmability of UEs

For the success of multivendor mobile networks that span over the whole globe, standardisation bodies such as 3GPP have played a significant role [DPS20]. Through this approach, various networking services and features have been developed in response to diverse use cases and challenging problems over the years. An example of a highly evolved area is air interface protocols [DPS20] that have been developed to combat various challenges and to provide a multitude of features for different use cases. Although this approach is proven to be vital in the success of multinational/multivendor mobile communications, the drawback is that it takes a considerable amount of time to introduce a new feature thus increasing the time to market and innovation. This is mainly the result of having to reach a consensus between operators, network and, device vendors who may all have different priorities. Those diverse and sometimes contradicting priorities often dictate adapting a simple-to-implement solution that may not be desirable in the long run. The limitations due to the long time to market new services and features is even more pressing in dedicated networks, where enterprises call for an integrated networking solution for their operation. An E2E architecture that facilitates innovation, adapts to new environments, and use cases, requires a UE programmability framework that complements the network programmability concept.

UE programmability (see Figure [3-5]) has the potential to be an enabler to help realising the vision of the "fit for purpose" promise of dedicated networks for both legacy and upcoming use cases in 6G. In a high-level view, a programmable UE enables introducing new features through APIs rather than having to standardise new interfaces. An E2E architecture encompassing programmable UEs will realise a truly adaptable network able to introduce new changes, enabling faster time to market, faster innovation, and support of verticals to name just a few.

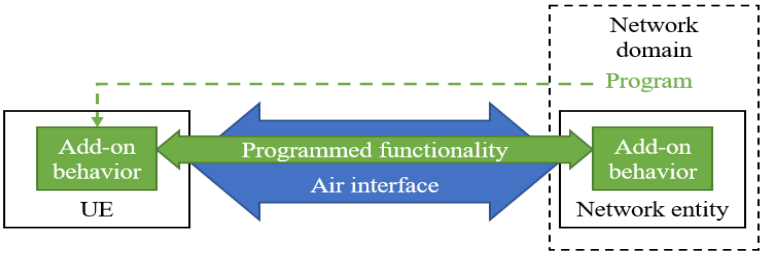

Figure 3-5: ***UE programmability [HEX21-D51]***

3.3.1.2 Programmable networks

The mentioned programmability of UEs is the first pillar for the realisation of an E2E programmable network. This implies the programmability and remote change of network functions and operations, also involving the modifications of devices' parameters. This will be pivotal to address the level of sustainability that 6G is expecting to achieve (more details on 6G sustainability topics can be found in Chapter 6). The E2E programmability will help reducing the costs of upgrading network infrastructure's potentials to new verticals that may come from the society. It is important to highlight that network programmability will involve all network resources and devices, enabling the adaptive management of the network continuum (see also Section 3.6).

This idea can be seen as the point of arrival of a concept that has been investigated for a couple of decades, namely, the Wireless Network Operating System (WNOS) [GBD+18]. This implies that a software network abstraction, consisting of all the microservices and agents implementing all the

network services, functions, and operations, can ensure automated network management. This can be centralised and/or distributed to deal with issues such as geographically localised management and operational subtasks of specific microservice' or chain of agents.

However, it is important to keep in mind that 6G aims for ubiquitous connectivity and extremely low latency, together with massive interoperability of heterogeneous technologies and devices. This will involve the overall control and data planes of the network. To this end, the emerging data plane programming language P4 [BDG+14, BR18], with its platform-independency feature, can be a good candidate to program more general data plane devices, e.g., SmartNICs, NetFPGAs, ASICs, or even software, in different locations. P4 compiler needs to abstract different hardware devices and translates P4 constructs into device-specific configurations.

Finally, it is important to notice that applications running in the central cloud will also involve functionalities of other network areas such as base stations and extreme edge/edge clouds, as well as the network transport in between. In this case, the important perspective of programmability is the one that also refers to hardware acceleration. This would be beneficial for a cloud-native deployment, where the programmable devices experience centralised management together with bare-metal servers. The discussion above raises an interesting link between programmability and network intelligence towards network programmability automation.

3.3.2 Network automation

Already with 5G, network automation is increasingly required to manage the complexities of modern, dynamic, and heterogeneous networks, and provide the means to orchestrate services through efficient and effective coordination and decision logic. Following this development, 6G networks and services are expected to be increasingly complex, with the growing integration of a heterogenous network of networks and capability to distribute functions and services towards the edge and extreme edge. Therefore, with 6G, the main aim of network automation is to enable the concept of "zero-touch" networks management (see also Section 3.6.1). In practical terms, this translates into reducing human errors in network management and operations, reducing service provisioning time while improving time-to-market, and reducing network and security issues through closed-loop network operations. For this, AI and ML techniques will be key to achieving a high degree of automation in 6G networks, unlocking the potential of data analytics to assist network orchestration and operations.

3.3.2.1 Zero-touch network automation through AI

The foreseen increasing complexity of operating and managing beyond 5G/6G networks motivates the use of closed-loop automation of network and service management operations complemented with AI/ML. The ultimate automation target is to enable largely autonomous networks that can self-adjust to new environments and new use cases.

In this context, AI and ML are more than ever considered as key enablers for full automation and optimization in 6G networks. In particular, with respect to conventional (non-AI/ML) algorithms currently used for network and service optimisation and automation, the use of AI/ML techniques enables to learn from data and improves the ability to execute specific network management and operation tasks at scale. Moreover, with the increase in operational complexities due to the very heterogeneous nature of network technologies and domains, the effective management and operation of connections, devices, and services become a critical challenge. Compared to conventional approaches, AI/ML enables to handle, manage, optimise, monitor and troubleshoot multi-technology and multi-vendor networks, moving from reactive problem solving to proactive operations and learning. In practice, AI/ML can help in the migration of traditional network and service operations towards automation and intelligent operations.

However, their integration brings new technological challenges [BT20]. While the use of AI/ML is already happening to support management processes for 5G and B5G networks and services, their seamless integration with network and service orchestration platforms needs to improve significantly to be able to address the complexity and heterogeneity of 6G networks. First, 6G network automation

can rely on AI/ML techniques to assist network management operations at different levels, including network planning and network functions placement, resource scaling, arbitration and sharing, Service Level Agreement (SLA) management. Second, as part of the 6G E2E architecture, AI/ML agents and pipelines need to become more seamlessly and tightly integrated with the network and service orchestration platforms. This is because current existing interactions are mostly happening by embedding pre-trained algorithms within the network management decision logics to assist specific operations among those listed above [BME+20]. In practice, the interaction with AI/ML orchestration engines needs to leverage on common services for algorithms and agents Life-Cycle Management (LCM), through unified APIs, data models and, metadata for capabilities discovery, data source requirements, as well as for the execution of algorithm training (and re-training), serving and evaluation services.

Therefore, in light of the 6G E2E architecture, the envisaged zero-touch network automation requires the implementation of a management framework designed for closed-loop automation and optimised for data-driven AI algorithms. In this context, a set of relevant standardisation efforts are working on network automation and are defining the required technical enablers to implement it. Among these, it is worth to mentioning the Zero-touch network and Service Management (ZSM) framework [ETS19a], and the TM Forum Autonomous Networks Industry Standards [TMF]. Beyond ZSM, deep integration and cooperation of per-domain network management and orchestration engines (e.g., responsible for managing network and resources at the heterogeneous extreme edge, edge and, core domains) have to be combined with local AI/ML orchestration engines to significantly simplify and ease the E2E 6G network automation processes. In particular, this can be enabled by considering abstractions within the network and between network layers as key in the enhancements and implementation of 6G network automation. Indeed, networks and network edges are currently migrating towards an ever-growing variety of heterogeneous network functions and applications, bringing to increased complexity in network management routines. Therefore, technology domain abstractions and separation of concerns in multi-domain and multi-vendor environments are fundamental in the path towards 6G network automation and programmability, where the E2E network management has to be implemented through highly cooperative local per-domain functions. The ongoing softwarisation trend already enables new levels of automation in network and service management, and in 6G this needs to make a further step to fully integrate several management loops and steps across different domains, including monitoring, analysis, fault management and orchestration for self-adaptive 6G networks.

3.3.2.2 6G analytics as enabler for network automation

Data analytics applied to network management lays its foundations on the capability to collect and process data potentially coming from multiple sources. This allows to dynamically elaborate the most efficient strategy and automatically re-configure and re-adapt the network and the services running on top of it. Already with 5G, the network analytics capabilities are being continuously enhanced over the last 3GPP releases with the aim of integrating AI and ML techniques to provide more accurate analytics reports (i.e., statistics and predictions) [23.791]. However, these analytics functionalities are still mostly considered as a closed box, limited to the boundaries of single operators.

From a practical perspective, a general framework for data analytics and network automation is defined (and continuously enhanced) for 5G networks. In particular, a dedicated network function, Network Data Analytics Function (NWDAF), is defined within the 5G core architecture. It provides data analytics services [23.288], mostly for mobility management, QoS provisioning and adjustment, policy configuration, SLA guarantee assurance, allocation of edge resources, and load balancing. For this purpose, the NWDAF may collect data from other 5G core NFs, other NWDAF instances, the Operations Administrations and Management (OAM), or from an Application Function (AF). Similarly, a Management Data Analytics Function (MDAF) is introduced in the 5G network management architecture to provide data analytics services in support of decision logics of other functions [28.533], such as network slice and network slice subnet management functions. While for 5G these analytics functions rely on pre-defined interfaces to collect and process data related to a limited set of source types, with 6G this will evolve to a more dynamic and open approach where data can potentially come

from any entity in multi-domain and multi-operator scenarios, and analytics capabilities should evolve towards the support of any requirement coming from any 3rd party application.

In addition, the 6G architecture has to consider new deployment models for these network analytics functionalities, evolving the current 3GPP work on hierarchical and distributed analytics [23.288], towards a flexible architecture with cooperative local per-domain analytics functions that provide AI and ML capabilities and share their outputs and outcomes. The goal is to uncover hidden trends and derive insights on the 6G network behaviour and network performance in a more automated fashion. This can be supported through AI-based anomaly detection solutions that learn the pattern of network behaviours by monitoring and analysing heterogeneous network data, including knowledge gathered from multiple domains and analysed holistically to derive cross-correlations.

In summary, AI and ML techniques will provide the required analytics capabilities to reconfigure the network for resolving or preventing anomalies by exploiting model-driven network programmability, and therefore deliver advanced 6G network automation solutions. Moreover, predictive network and service analysis will be able to output rich recommendations for supporting network operation teams decisions in those cases where human validation will still be required.

3.3.2.3 Network of networks automation

The concept of the network of networks was first realised with the Internet, which was interconnecting different kinds of computer networks of different sizes. Nevertheless, the concept of the network of networks has extended its meaning with the advent of cloud/edge computing, 5G and, upcoming 6G. Now, the whole E2E continuum of the communication network is a network of networks, which embraces not only the RANs, the operators' edges and cores, and the Internet, but also the IoT and the cloud/edge data centres, which are networks themselves. Moreover, in the context of 6G and the Tactile Internet [FLS+21], the continuum network of networks will also consider the intra-body networks, combining intra-body sensors and molecular communication networks, where for example biological molecules are used for the communication. Networks will contain an ever-growing variety of functionality and services that, coupled with customised devices, lead to difficulty in network management and service provisioning with traditional means. Flexible abstraction and separation in a multi-vendor environment strongly rely on Network Function Virtualisation (NFV), which also is the key enabler of programmability. Then, NFV and programmability are pivotal in the move towards network automation.

The softwarisation trend enables new levels of automation in network and service management and orchestration, as well as full closed-loop automation (see also Section 3.6.1 for more details on management and orchestration of different administrative domains). The proposed Hexa-X architecture principles and enablers target a fully softwarised, programmable, and intelligent 6G network, for a further step towards the automation of the network of networks. In fact, in future networks, the automation will not only embrace individual functions or operations (or their sub-modules), but it will also involve the whole architecture. The vision of Autonomic Mobile Virtual Network Operator (AMVNO) [GRA18a] [GRA18] represents the complete automation of the E2E communication networks at each level, from the physical network infrastructure, consisting of deployable access points e.g., base stations carried by Unmanned Aerial Vehicles (UAVs), SDN switches, datacentres of different sizes, and satellites counterparts and High-Altitude Platforms (HAPs). This dynamic architecture is enabled by the softwarisation and programmability not only of the networking and functionalities of the communication network but also of the protocol stack, the so-called Programmable Protocol Stack (PPS). The key element of the AMVNO is the autonomic manager, which contains the intelligent hypervisor and the intelligent business hypervisor, which are respectively focused on the network aspects and business-related aspects. In fact, the latter can use AI to adapt economical aspects, such as pricing and expenses for network expansion. The idea of mobile operators using AI for smarter capital spending, automation, and simplification in the back office, predictive analytics in marketing and sales, more efficient customer retention and support has already been suggested in 2017 [GRA18a].

3.3.3 AI and AI as a service

As detailed in [HEX21-D12], an AI service following the AIaaS concept (see Subsection 3.3.3.1) can be consumed either by NFs or by AFs external to the network (e.g., via network exposure) by submitting requests for ML-based inference decisions to the AI service. Architecturally, several network entities, and their corresponding (standardised) interfaces, would be needed for the efficient offering of AIaaS in an open manner, across systems owned and managed by different network operators and vendors, as described in the following Table 3-1:

Table 3-1: ***Functional entities supporting AIaaS operation***

Functional entity	**Duties**
AI orchestration function	Responsible for maintaining an overall view of available in-network AI resources and topologies (e.g., learning federations), adding/ removing instances of data analytic functions and, AI agents.
AI repository function	Responsible for the registration of available AI agents and their offered services.
AI policy enforcer	Responsible for implementing the recommended learning/inference policy. For example, learning data gathering and further pre-processing may be decided in case the AI success monitoring function flags an experienced degradation in inference accuracy when given AI agents are selected to provide such output to a calling entity (e.g., client application).
AI success monitoring function (optional)	Responsible for monitoring the quality, efficiency, and security of the implemented policy. Confidentiality breaches could be also monitored, however, (near) real-time logs of such breaches are quite challenging.

A new AI-enabled architecture aims to support distributed AI services, needed for supporting AI as close as possible to the application, AI service chaining (in the sense of assisting with AI traffic flow between AI services in the network) needed to accomplish specific AI tasks, as well as cross-domain AI service consumers and data producers. The in-network AI architecture, as proposed in [HEX21-D51], also aims to support AI-enabled access control considering attributes such as user, data object and environment information, efficient transfer of large amounts of data, network and application-specific analytics, and the sharing of AI models, once available and updated. The new architecture supporting AIaaS will be employed for enabling different learning services, such as Federated Learning (FL) and Explainable AI (XAI). Once a consumer requests the FL service, mechanisms to allocate resources and instantiate the required functions are needed. Based on the service requirements and mobile device's capabilities, the AIaaS-supporting 6G network will be able to decide which functions of the service will be instantiated. On one hand, the mobile device may produce data, receive the global model from the FL server, build local AI models, and make decisions based on it. On the other hand, resource-constrained devices (e.g., sensors and actuators) may delegate data aggregation in a centralised manner aiming to form a traffic environment's digital twin instantiated at the edge of the network (e.g., in the form of a Multi-access Edge Computing (MEC) application), while the device only takes care of

sensing the environment. In that case, inference-based decisions and/or predictions can be taken by an AI agent using the formed digital twin as one of needed the inputs. Also, the FL application server may be implemented as one or more functions residing in either the cloud or the edge environment, based on both the service requirements and edge nodes' capabilities. In the context of MEC, MEC federation [ETS22] may be employed to extend the scale and coverage of an FL service, by allocating functions under the domain of several MNOs. Based on how the different FL and XAI functions are distributed in the device-edge continuum, flexible protocols will be envisaged, accounting for the large dynamics in, e.g., radio environment changes affecting connectivity and, in its turn, the ability of a device to attach to/detach from a learning structure (e.g., FL setup).

3.3.3.1 AIaaS – Bringing network knowledge to the edge

An incurred challenge of bringing intelligence to edge network devices is how to design a protocol allowing a UE carrying a local ML model (obtained by e.g., training a Neural Network (NN)) to seamlessly exploit the knowledge of large parts of the network, useful to its locally undertaken inferencing tasks, by attaching to/detaching from different learning deployments (e.g., federations) across multiple operator areas. Criteria calling for such attachments/detachments are (i) UE and coverage-providing network node mobility, (ii) unavailability of an AI agent (e.g., FL aggregator due to e.g., a detected security attack), (iii) low-quality connectivity (impacting crowdsourcing of local models), (iv) increased model aggregation latency and others.

To address such challenges, the interaction between the inference requesting entity (e.g., an application instantiated at the UE or machine, such as a robot) calling for inference-related support and the available AI agent(s) carrying relevant model(s) needs to be facilitated. This can take place either directly, in case such direct interface exists, and the AI agent has been pre-discovered and selected, or via an AI Information Service (AIS) and its corresponding AI Application Programming Interface (AI API) that can be introduced to a 6G network and implemented over an interoperable network interface. In the second case, an AIS consumer may communicate to the AIS input information relating to a user/client application-specific task calling for an inference-based recommendation (e.g., road object identification when on the road) and performance requirements (KPIs) relating to e.g., inferencing accuracy, energy efficiency, E2E delay, inferencing data integrity, and others. All these criteria are filtering criteria for AI agent selection. Then, the most appropriate AI agent will be selected to either provide a copy of its contained model to the requestor or directly provide inference output. The approach involving an AIS acting as a "mediator" between an AI consumer (e.g., a UE) and an AI capability provider (AI agent), is illustrated in Figure 3-5.
As a result, considering each selected AI agent, the UE will be able to share its local model updates to the AI agent(s) it is subscribed to and obtain learning system parameter updates by the subscribed AI agent(s) or direct inference output for, e.g., infrequent requests. The advantage of the proposed AI-supporting protocol (at the application layer) is to enable an AIS consumer (e.g., end user device) to exploit knowledge available across the network without violating the privacy of data or model contributors. The AIS consumer receives a trained (and, therefore, "plug-and-play") model, or access to using such a trained model in the network, derived through network internal processing of the available knowledge. Such available knowledge refers to relevant training data that were gathered based on an AI agent "skillset" criterion. For example, a model issuing QoS predictions for vehicles will be trained via a sensor, camera, etc. measurements received from given locations of interest and not based on application data produced outside a road environment. Learning data routing to the appropriate AI agent is part of ongoing research.

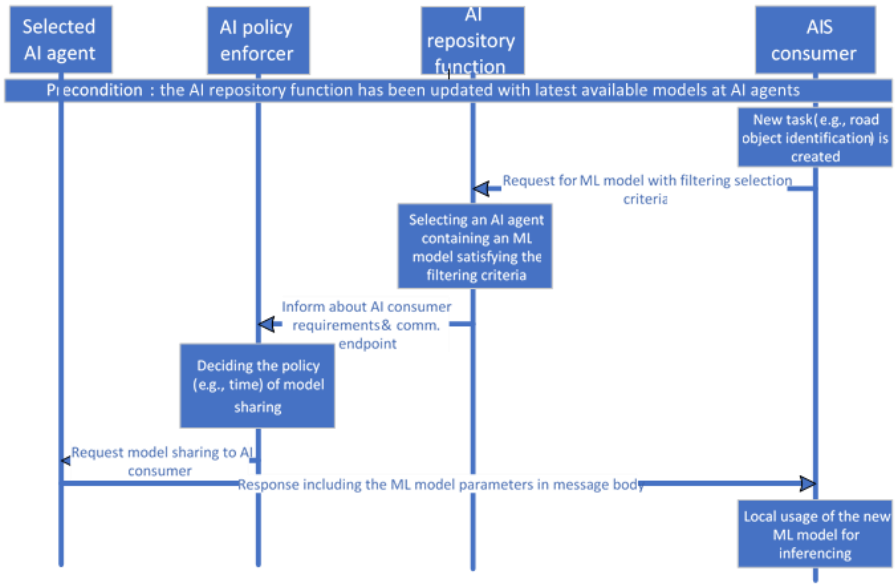

Figure 3-6: ***Signalling flow for requesting/delivering of new ML model satisfying AI agent selection criteria posed by an AI consumer (e.g., UE)***

3.3.3.2 Protocols for AI

On the one hand, the 6G network should be adaptable to dynamically respond to variations in mobile service demand, as well as changes at the infrastructure level. On the other hand, large-scale network reconfiguration and real-time management could result in high costs, making unsupervised management a viable option. 6G is also predicted to allow for large-scale AI/ML agent deployment. In factory automation, for example, the control process will be designed to be AI-driven. The control operation will be depending on the operation of traditional sensors, actuators as well as AI-based ones.

To realise effective i) AI/ML for communication and ii) communication for AI/ML, new functional entities and protocol-level aspects are necessary. However, it is still unclear which requirements, signalling, and other aspects of their realisation over the air interface would be required for AI models, including the transfer of AI-related data, models, and algorithms between gNB and UEs. For training, inference, and maintenance, protocols are necessary. As evidenced by recent breakthroughs [GoogleAI], training methods and algorithms are rapidly evolving. As a result, the protocol aspects addressing the training of such models, such as model/data distribution techniques, loss minimisation algorithms, compression techniques, and so on, should be adaptable to the rate of evolution and innovation. Inference and actuation requirements will be evaluated to determine what needs to be augmented in the protocols (e.g., signalling) that address the input/output from AI/ML agents.

The seamless transfer of learning between domains and planes, as well as the exchange of data analytics across distributed network microservices and agents, induce new problems in terms of developing and implementing efficient and effective protocols. To begin with, effective use of AI methods necessitates the design of novel, reliable and efficient protocols for data mining and traffic engineering. These protocols are essential for virtual intelligent functions/agents/services to get actual network status and, as a result, make decisions and take actions on the existing network infrastructure. Second, collaborative-distributed intelligence will be performed by multi-agent systems. Effective synchronisation and data sharing mechanisms will be required for these collaborative AI techniques. It is vital that these intelligent network entities will execute on many servers or data centres throughout the network infrastructure. As a result, reliable communication mechanisms for exchanging analytics and their processed results throughout the multi-agent collaborative system are required. Packet-based protocols as Precision Time Protocol [IEEE19] are particularly efficient and adaptable in case of synchronisation. They may, however, have communication challenges, resulting in synchronisation

errors. The development of more complex synchronisation methods for the design of novel, reliable and efficient protocols and collaborative AI techniques described above is a significant task.

A relevant example of collaborative AI is Federated Learning (FL) since it allows involved entities to share ML models while preserving data privacy. In order to allow mobile devices to leverage FL services, they will need to query a registry of available federations in a given coverage area, each providing at least information about its objective (e.g., optimisation of QoS for Vehicle-to-everything (V2X)) and requirements (e.g., minimum storage required for AI models on the mobile device). Protocols for allowing mobile devices to join/leave a federation or even trigger the creation of a new one will also be needed. Besides receiving the updated global model from the FL server, mobile devices will need to effectively transfer local ML models or aggregated data to the FL server. Considering the expected huge number of mobile devices connected to the 6G network, employed protocols for the above operations must also be designed to be scalable by taking into consideration context information like federation type and scale, radio access conditions, and availability of computing resources. Moreover, when using XAI models to improve trustworthiness, the 6G network will need protocols to collect explanations about the inferences made by the algorithms themselves, by either explicitly requesting (i.e., using a client-server paradigm) or subscribing to explanation updates (i.e., using a publish-subscribe paradigm), balancing possible trade-offs between effectiveness and overhead of the different approaches.

3.3.4 Dynamic Function Placement

NF migration and placement problem has been studied in a number of previous works, such as [LAG19] and [HJS17]. In Hexa-X we extend this work to 6G architecture to consider Dynamic Function Placement (DFP) to a full fabric of different domains of the architecture, from the end-user domain up to the central cloud to enable continuum orchestration (Subsection 3.6.1). This implies that DFP needs to operate across domain boundaries of collaborating clouds. The domains need to expose their shared resources and relevant APIs for service discovery. Furthermore, the connectivity between the dispatched NFs instantiated in different clouds need to be set up in an efficient technology-agnostic manner to ensure full interoperability via network service mesh.

3.3.4.1 DFP for E2E architecture

DFP is responsible for ensuring optimal function deployments to provide differentiated services in a single domain, multi-domain, and cloud environments, as explained in Section 3.6 [HEX21-D51]. For the functions with enabled DFP features, monitoring of the specified cross-layer KPIs must be supported. The monitoring info is used in decision making where the performance of the current functions is evaluated, and a need for changing function instance numbers or locations is determined. Decision-making could be based on the usage of AI/ML techniques, for instance, provided by AI/ML services for Orchestration and Management functionality as shown in Figure 3-1.

In one sense, DFP could be seen as one of the core functionalities on top of what the orchestration framework's Life-Cycle Management (LCM) provides. The main responsibilities of DFP can be characterised as i) relocation of function instance, and ii) runtime context transfer for the instance relocation, which are closely related to more traditional LCM functionalities like replica management and service scaling up and down. Respectively, the separation of responsibilities and roles between LCM and DFP is not always clear and depends on the implementation aspects. For instance, the instance relocation raises some new technical challenges like how to define extractable and hence movable runtime state for any function and how to securely move such state potentially between domains. Additionally, regarding the scope of orchestration and management (incl. DFP) operations, special focus should be put on the notion of the domain, i.e., how to support multi-domain operations (See also Section 3.6.1).

For different layers in the layered E2E architecture, i.e., the application layer, the network service layer, and the infrastructure layer (in Figure 3-1), monitoring of layer specific KPIs and KVIs are required. Depending on the use case, monitoring could target different layers with specific KPIs/KVIs. For instance, for end-user services, the "health" of a service is generally measured in the application layer.

Respectively, the infrastructure layer provides resource-specific metrics to be used in operations done in the network service layer. There is a special relationship between the infrastructure and the network service layers visible in definite orchestration and management operations where the existence of special hardware resources in the infrastructure layer governs how the operations can be executed. In other words, functions can only be placed in physical nodes where the required hardware resources are available. In addition, the extreme edge represents the infrastructure layer part with the most limited computing resources and sets new challenges for the cloudification itself.

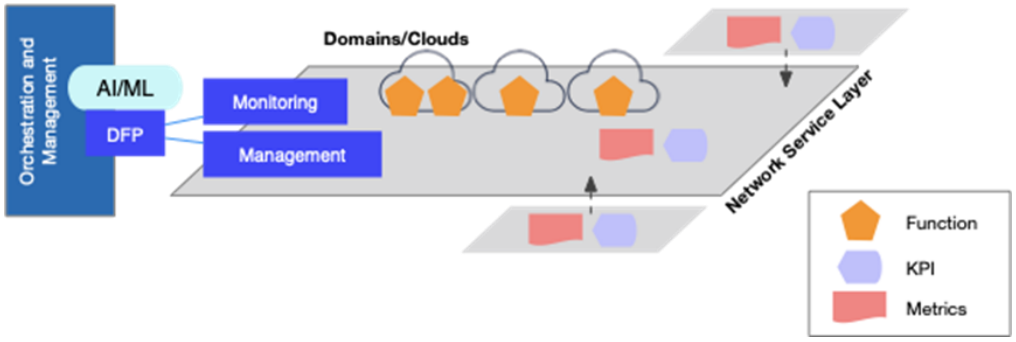

Figure 3-7: ***DFP layered view***

Figure 3-7 depicts how DFP is positioned in the layered E2E architecture and how cross-layer KPIs and metrics are required in the network service layer, which is the focus of the DFP management operations.

3.3.4.2 Network service meshes

To fulfil the flexibility and dynamicity requirements considered for 6G for cloud-native applications, services, and infrastructures, the Hexa-X research targets to introduce and develop the concept of network service meshes to support connectivity between the Network Functions (NFs). The different flavours of edge computing are expected to become more and more relevant for 6G architecture and services as more latency sensitive applications are introduced (e.g., V2X, factory automation, and Extended Reality (XR)). Hence, network service meshes will need to interconnect a large number of application-to-application and function-to-function communication assets over flexible mesh topologies between the edge clouds and the central cloud(s). From a practical perspective, network service meshes will provide automated virtual network connectivity services with advanced User Plane (UP) and Control Plane (CP) functionalities. The aim is to support dynamic, flexible, and transparent Layer-2/Layer-3 connectivity among the various cloud-native applications and functions. On the one hand, this requires mechanisms and capabilities to forward and route data among the involved virtual network service endpoints according to the cloud-native application and related mesh topology requirements. On the other hand, network service meshes require proper control and management functionalities to provide discovery, routing, and connection management capabilities for virtual network services exposure.

With respect to currently available technologies and solutions mostly related to traditional NFV deployments, the main goal is to create a flexible and responsive network mesh environment that could be more suitable for cloud-native applications and functions. Indeed, there is a need to evolve the mostly static NFV-like forwarding graphs with more agile communication patterns. In this context, similarities with the 5G SBA can be exploited (e.g., in terms of service discovery and service registry) to adopt publish-subscribe and request-response mechanisms in both asynchronous and synchronous network service mesh communications. For this, the control and management of the network service mesh need to be flexible to allow for dynamic cloud-native applications and functions that will adapt to 6G service mobility and elasticity of different use cases.

3.4 Enablers for flexible network

Flexible networks intend to enable extreme performance and global service coverage. The network functionality and architecture must then be flexible enough so that it can adapt to different topologies.

3.4.1 New mobility solutions for flexible network deployments

The deployment of mobile networks has become increasingly complex and diverse with every new generation. During the 4G standardisation there were many discussions about so-called Heterogeneous Network (HetNet) solutions, i.e., how networks with both wide-area macro and small-cell pico base stations should cooperate. The extension of the radio spectrum into mmWave in 5G added yet another aspect to flexible deployment. 6G deployments will include nodes using even higher sub-THz spectrum (e.g., in the 100-300 GHz frequency range) with limited coverage as well as nodes at low frequencies with seamless coverage, as illustrated in Figure 3-8. Furthermore, the number of network solutions for capacity and coverage is also expected to increase in the 6G timeframe. These include solutions such as Distributed MIMO (D-MIMO) networks, Non-Terrestrial Networks (NTN), campus networks, mesh networks, and cloudification of the network elements. Thus, 6G will be a network of networks.

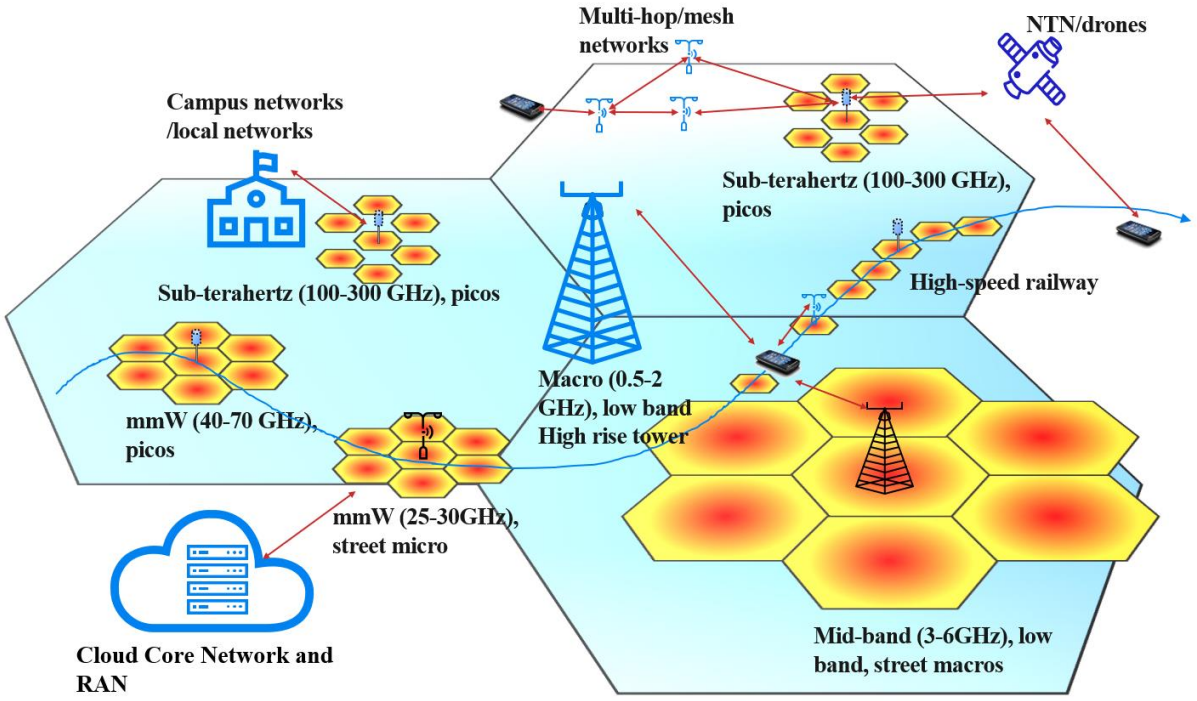

Figure 3-8: ***The 6G network of networks***

As mobile broadband is becoming increasingly critical to society, the architecture of 6G must support reliability and resilience beyond 5G, both in terms of service and infrastructure provisioning, when connecting through any of the diverse connectivity options.

To cope with the expected limited coverage of the sub-THz nodes and the expected higher reliability of a 6G (compared to 5G), a new more flexible Multi-Connectivity (MC) solution can be one possible solution (see also Chapter 5). The new MC solutions can be built on either Dual-Connectivity (DC) or a Carrier Aggregation (CA)-type of solution. Further on, to handle a more flexible network, the new MC solution should include more than two connections e.g., one master node and two secondary nodes, and the connections should be decoupled in DL and UL to allow, e.g., three connections in DL but only one in UL. A general disadvantage with the DC solutions for New Radio (NR) is the implementation complexity of the 3GPP specification, for example, the numerous architecture options for DC between

LTE and NR and the message exchange over the Xn interface between gNBs [38.331] so care needs to be taken to reduce complexity. Other ways to ensure reliable mobility may be to use mesh networks, where, for instance, Integrated Access and Backhaul (IAB) nodes can create a dense network without the need for wireline fronthaul and backhaul.

With the combination of terrestrial networks and NTNs it will be possible to achieve 100% global coverage, including the oceans [BFC21]. NTNs can likely provide a lower capacity per km² than terrestrial networks, but at a reasonable cost. Thus, an NTN is suitable for rural areas with low population density. For urban areas, there will always be a need for terrestrial networks. There are two types of architecture options for NTN: Transparent and Regenerative payload architecture. Transparent is the simplest type, where the NTN basically serves as a relay of the signal between the UE and the base station on the ground. The regenerative architecture is equivalent to having the base station (RAN) functions onboard the satellite. The main research question for 6G is how the NTN and terrestrial network mobility will be solved. Since the Low Earth Orbit (LEO) and Medium Earth Orbit (MEO) satellites move, it may be necessary to find solutions that minimize the number of handovers and the signalling needs for mobility robustness. Another important research topic for 6G NTN is the actual architecture solution, e.g., if regenerative or transparent or a hybrid split should be used [BGS+20, BFC21].

D-MIMO may be a component for 6G systems, due to the potential improvements in spectral efficiency (see also Chapter 5). With D-MIMO, is possible to utilize L1/2 mobility. The L1/2-mobility system relies on a D-MIMO deployment with several Access Points (APs) connected to a central unit via a high-capacity fronthaul transport network. In a given region, all the APs connected to the particular CU typically utilise the same resources but without fixed cell borders, which is also referred to as a MIMO cluster area.

3.4.2 Campus network

A campus network is a network made up of an interconnection of Local Area Network (LANs) within a limited geographical area. Campus networks have been important mainly for companies for years [FGS20]. Non-Public Networks (NPN) are a specific instance of campus networks. In fact, they are LANs or local combination of LANs, which were originally specified for industrial customers [BFC21]. So, why do we need to talk about new campus networks in 6G? As discussed in the sections above, 6G will be a fully softwarised, flexible, and micro-service-based architecture. Further, 6G will massively employ in-network intelligence and multi-agent systems for network automation, and it will require complete trustworthiness. These aspects come from the support that the network will provide for very sensitive verticals related to the Tactile Internet such as eHealth and Industry 4.0. Such verticals require very demanding concurrent KPIs in terms of latency, reliability, and resilience. Moreover, they need high level of security and privacy. These are the motivations behind the prominent role that new campus networks will have in 6G. However, such networks require to be enhanced with 6G capabilities.

Industry 4.0 will require massive robots' automation and collaboration, with other machines and humans. Robots, sensors, and other hardware will need Ultra Reliable Low Latency Communication (URLLC) of 6G. Moreover, data belonging to humans, machines, and the campus network will be securely stored and computed within the edge nano data centre, within the campus network. In fact, in order to ensure low-latency, small data centres are placed within the campus network so that no external data centre is used. The increasing number of interconnected humans and 'things' in URLLC-related verticals will imply continuous heterogeneous data mining, classification, and transmission to a dedicated intelligent edge data centre within the campus network. Intelligence will not only be applied for network management and operations but also for robots training, acting, and deciding in order to better work autonomously or to assist humans. In this sense, 6G campus networks will represent specialised networks of LANs, also providing edge computing resources for massive data handling and communications.

Furthermore, the 6G scenario will also be three-dimensional. This means that the terrestrial network will seamlessly interwork with aerial platforms and satellites in a unique network architecture. Then, another novel aspect will be the three-dimensionality of campus networks, where the RAN might be

based on UAVs. On-demand mobile base stations based on UAVs can provide low-latency, and secure and resilient RAN, however, the MEC and the fronthaul/backhaul might be realised via HAPs [GCZ+20] [BGS+20]. Both HAPs and UAVs are equipped with communication interfaces and computing/storage. In this way, in-HAP MEC can also provide computing resources and virtualisation means to define E2E resource management and virtual network slicing. Additionally, it can also host value-adding service for network and users' applications.

With respect to terrestrial campus networks, in three-dimensional campus networks, the networking operation and coverage assurance become critical. The HAPs have to adjust their altitude and position based on environmental and atmospheric conditions. This is very important to ensure constant service availability and satisfaction of KPIs requested by the verticals hosted. The HAPs have longer battery life than UAVs and can potentially carry heavier weight and cover a relatively larger area [GCZ+20].

3.4.3 Mesh and device-to-device integration

Device-to-device (D2D) communications enable devices to communicate directly in an infrastructure-lean manner, offering significant gains. First, the proximity of users can allow even higher throughput and lower transmission powers. Moreover, radio resources can be reused more efficiently. In addition, D2D communications can enhance the coverage and capacity of cellular networks through UE relaying. The “Network of Networks” approach considers the usage of technologies for supporting flexible topologies to increase the availability and reliability of the connection. In order to support this, new intelligence will be needed to make decisions “on the fly”. Cost efficiency (e.g., limited resource and energy consumption) should underpin all operations.

Hexa-X will investigate two main approaches for D2D communications:

1) Distributed approach, where devices use mesh protocols to find neighbours and connect autonomously.
2) Centralized approach (or operator controlled) in which devices will connect to each other as designated by a central management entity (e.g., an operator).

The following architectural challenges are relevant to both distributed and centralized D2D communication:

- Defining the trust level for devices participating in the D2D/mesh network.
- Unified modelling of nodes and devices, in terms of network and computational resource characteristics, capabilities and constraints.
- Definition of interfaces to control and interact with devices for resource advertisements, synchronisation, reachability verification, etc.
- Selection of best possible nodes and devices depending on specific parameters (e.g., position, signal quality, battery level, availability, reachability, available computational resources, etc.).
- Integration with network and service orchestration for seamless management, control, and enforcement of D2D/mesh network communications.
- Methods and procedures for discovery of nodes and devices (including synchronisation aspects for capabilities advertisement).

3.4.4 Edge cloud integration

The Hexa-X architecture will also develop and validate the edge-to-network integration concept enabling a unified control plane, integrated network orchestration methods, and intelligent resource and network function placement and management across the continuum from edge to the operator’s networks.

One of the major proposals to feed Hexa-X architecture enablers is to develop flexible strategies for the deployment of edge-enabled network functions to support different local and time-sensitive services as well as coverage extension [HEX21-D51].

The different devices and computing/storage resources in the local area will communicate with each other and with the rest of the network to provide seamless services in a highly dynamic environment with varying traffic and changing channel and mobility conditions. The network-integrated edge will enable intelligent decisions on computing and storage resource use, including user device cooperation and cooperation with MEC. The edge-to-network integration enables support of multi-domain end-to-end network slicing and network management. It is necessary to extend the slicing paradigm onto all edge resources so that the slice deployment is supported in an optimised way over all possible resources with changing environments in an E2E manner especially the edge-enabled ones allowing them to meet the requirements of URLLC-like services. AI and ML techniques are also used for integrated edge-enabled service slice definition and selection.

The main security concerns are related to the level of trustworthiness of the integrated edge components or resources with regards to the E2E security, so that all network components are managed by trusted entities. The edge paradigm extends the computing and storage towards the edge of the network including even the user devices. Thus, security constraints in a network-integrated or "integratable" edge are mainly related to the extent that the authentication of edge nodes, the existence of a trust model ensuring the trustworthiness of the integrated edge components or resources and also a secure and privacy preserving distributed user data storage in the light of the latest European legislation about privacy preserving (General Data Protection Regulation - GDPR) can be guaranteed.

3.5 Enablers for efficient network

A new smart connectivity platform capable of fulfilling 6G requirements will require an architecture that can integrate legacy networks and mission-critical networks while increasing dependability, coverage, and reliability. Important research activities to investigate and develop for 6G include, for instance, determining the right level of modularisation, focusing standardisation on the basic functionality while ensuring that it is possible to build more advanced functionality on top, e.g., using network programmability (see also Section 3.3.1). Integrating and leveraging such architectural trends from the onset of the 6G concept development is expected to lead to increased flexibility and cost efficiency, as well as decreased overall complexity.

With 5G, 3GPP enabled a cloud-friendly 5G core network, where the so-called Service Based Architecture (SBA) is based on the principles of network services offered by network functions communicating via web-based APIs. The cloudification trend is expected to continue for 6G enabling novel network designs, e.g., cloud-optimised network procedures can be obtained considering NFs capable of accessing any (authorised) network information with limited (or no) nested interactions among NFs. The key challenge is to design a 6G architecture that can fully utilise and interact with the cloud platform with regards to speed of development and reuse of common cloud components, balancing the need to standardise business-critical interfaces with the fast evolution of IT tools, such as DevOps.

3.5.1 Architecture transformation with cloud and SBA

With a cloud-native approach, the RAN and CN architectures should be possible to streamline, i.e., reduce heterogeneity and hence some complexity. In cloud environments, complexity could arise with hierarchical interactions among NFs, multiple processing points for a certain message, duplication of functionalities among functions, etc. In 6G, hierarchical interactions might be reduced if, for instance, a network entity/NF could access any network service or relevant network data, without the need of relying on intermediate entities/NFs. The number of processing points might be reduced by redesigning network functionalities with the aim to perform all relevant processing for a certain network task in a single point. This might result in a different placement of services exposed by the NFs, as well as in a possible reduction in the number of NFs. As a consequence, there would probably be fewer business-critical interfaces.

One example of a streamlining opportunity is related to signalling. In 5G, signalling of a typical procedure involves the UE, gNB, and usually several NFs in the CN. Figure 3-9 shows an example of how a particular action involves several interactions between different NFs, all of them activated in sequence since the current network architecture is hierarchical. This action [23.502] could, e.g., be a handover request and handover execution. Involved NFs would in that case be Access and Mobility Management Function (AMF), Session Management Function (SMF), and the User Plane Function (UPF).

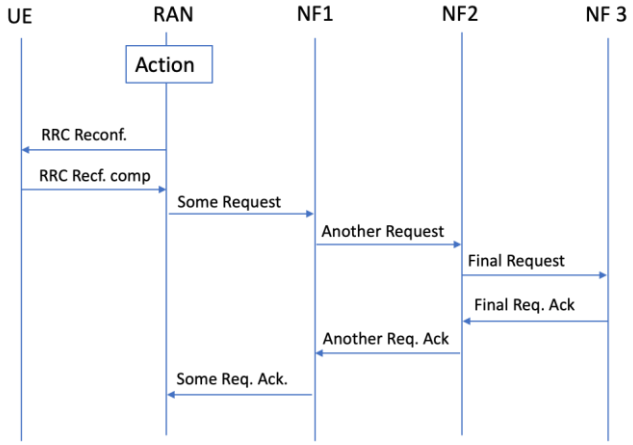

Figure 3-9: ***Principle of signalling flow in current networks***

With a cloud-native design of the CN and RAN, it should be possible to perform the requests to NFs in parallel and, hence, optimise the overall signalling speed. Figure 3-10 shows an example of how the particular action can be simplified in a cloud-optimised network architecture without the hierarchy between NFs. In this case, RAN could directly notify the different NFs about the event. Overall, the example in Figure 3-10 leads to more efficient use of signalling resources and is future proof (if other NFs need to be added).

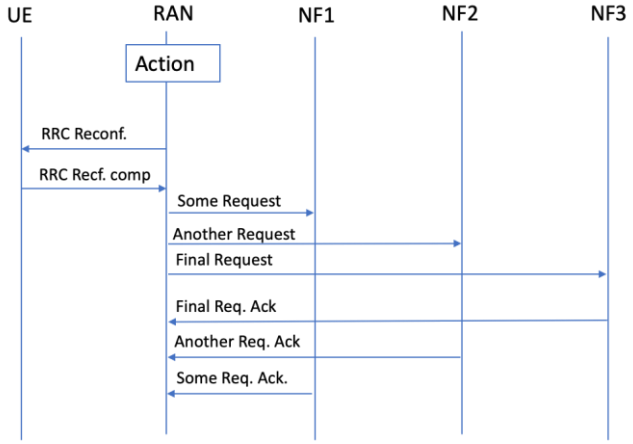

Figure 3-10: ***Simplified signalling for the same action in a future network***

3.5.2 Compute as a service

Compute-as-a-Service (CaaS) is a use-case-enabling service approach, as described in [HEX21-D12], which is aimed to / shall be used by any device (static stationary or mobile, IoT, handheld, etc.) or network infrastructure equipment that chooses to delegate demanding, resource-intensive processing tasks to other parts of the network. The network nodes for workload addressment/execution are chosen

as providing more powerful compute nodes, which are also of higher availability at the time of workload generation. These service-offering compute entities can be either onboard devices other than the requesting one, or, for example, integrated to edge cloud servers at the infrastructure side. In the CaaS case, external compute resources can be made available to a specific entity or user device through a well-defined open interface. The basic principles relate to an offload of processing tasks to external compute resources. In this context, some of the needed features to be defined, as part of a 6G network architecture design, are the following:

- A general interface providing access to external computational resources.
- Mechanisms for discovery/detection of available compute resources (e.g., via a general register reachable by the CaaS provider).
- A functional entity (e.g., central controller/workload orchestrator) that decides when to offload (fully or partly) a processing workload.

The decisions on whether to offload a processing workload, and, if so, where to delegate the workload, are based on the knowledge of currently available resources of network nodes (or prediction of future availabilities) and taking into account requirements relating to performance (e.g., the delay for producing workload output and dispatching it to the requestor), the energy footprint of the workload delegation and the trustworthiness of the network node(s) offering their compute resources for workload processing. AI capabilities of the network can be exploited to orchestrate the task workload delegations - more details are provided in Section 3.6.

A key challenge in the CaaS concept is to balance the following - often contradicting - objectives: on one hand, the user should be enabled to access computational resources, which often consist of heterogeneous elements and architectures (e.g., CPUs, GPUs, FPGAs, etc.) and which would, therefore, need to be utilised at the maximum possible level of efficiency; on the other hand, the user should ideally rely on generic and simple interfaces that are abstracted from the underlying hardware and software (such as operating systems, etc.) to the extent possible. As a way forward, per [HEX21-D51], the proposal is to offer various trade-offs between flexibility, simplicity, and efficiency to users to choose from, depending on the underlying scenario. Three CaaS approaches can be thought of:

- Applications customised to assist with the processing of a specific workload (e.g., signal processing, object recognition, etc.) can be pre-installed by a service provider.
- Full access to “raw” hardware and software resources is provided to the service calling entity (e.g., device).
- Compute Virtual Machine (CVM) based approach: as a compromise between the previous two approaches, the proposal is to use a “Compute Virtual Machine (CVM)” approach building on an extension of the “Radio Virtual Machine (RVM)” [ETS17a]. Per [ETS17a], an RVM is an abstract machine that supports reactive and concurrent executions. A RVM may be implemented as a controlled execution environment that allows the selection of a trade-off between the flexibility of baseband code development and required (re-)certification efforts.

Regarding the third approach, any application code is first processed by a front-end compilation step, which creates so-called *Configcodes* following the RVM requirements. In a second step, a back-end compilation is performed to map the application to a specific (heterogeneous) target platform consisting of a specific number of resources (such as CPUs, FPGAs, memory, etc.). The process is supported by a library, which provides optimised functionalities to the developers. The Virtual Machine approach mainly consists of Data Objects (DO), which are connected to Abstract Processing Elements (APEs) through an Abstract Switch Fabric (ASF). A control unit interconnects the building blocks, as required.

Further details on these approaches and the usefulness of AI/ML techniques to implement CaaS procedures can be found in [HEX21-D51].

3.6 Enablers for service management and orchestration

As shown in Figure 3-2 (E2E architecture overview) the Hexa-X Management and Orchestration (M&O) function is a common functionality impacting all layers of the architecture: from the infrastructure components up to the applications. This is because this functionality is intended to provide a complete and effective E2E orchestration of all the services that could be deployed across all the different architectural domains, including the MNO's own domains, and also, other domains beyond the MNO scope (i.e., other external networks such as public/private clouds, hyperscalers, vertical's networks, and others).

The architectural design of these M&O mechanisms is being specifically addressed in the context of WP6 (currently underway), where the "goal state" of this system has already been defined by means of a set of specific features [HEX21-D61].

This section is structured in two technical subsections, each of them addressing the areas of network continuum orchestration (Subsection 3.6.1) and AI-driven orchestration (Subsection 3.6.2). The alignment of the M&O functions and design with the E2E architecture introduced in this document is also analysed in Subsection 3.6.3.

3.6.1 Toward continuum orchestration

As mentioned, E2E seamless integration management makes reference to the network services orchestration considering the integration of the management of services or infrastructures of different network domains, from the end-user's domain up to the cloud. The goal is that, from the M&O system, the diversity of all the infrastructure elements in the different network domains are managed jointly and following a unified approach, as a single resource pool on which network services can be deployed and orchestrated in a very efficient way, also allowing a high degree of automation degree.

This cross-domain, E2E seamless integration reflects the "Device-Edge-Cloud Continuum Orchestration" concept defined in Hexa-X, which implies the integration of distributed infrastructures across the different network domains from the end-user's domain up to the cloud (cloud, RAN, edge, etc.), considering the diversity of all these domains as if it were a continuum, rather than different isolated silos.

The Continuum Management and Orchestration concept is one of the major innovations in Hexa-X, since it expands the management and orchestration focus, when compared to the previous mobile communications generations (see Figure 3-11). The full development of this concept constitutes one of the main objectives in WP6, to be demonstrated during the project activities.

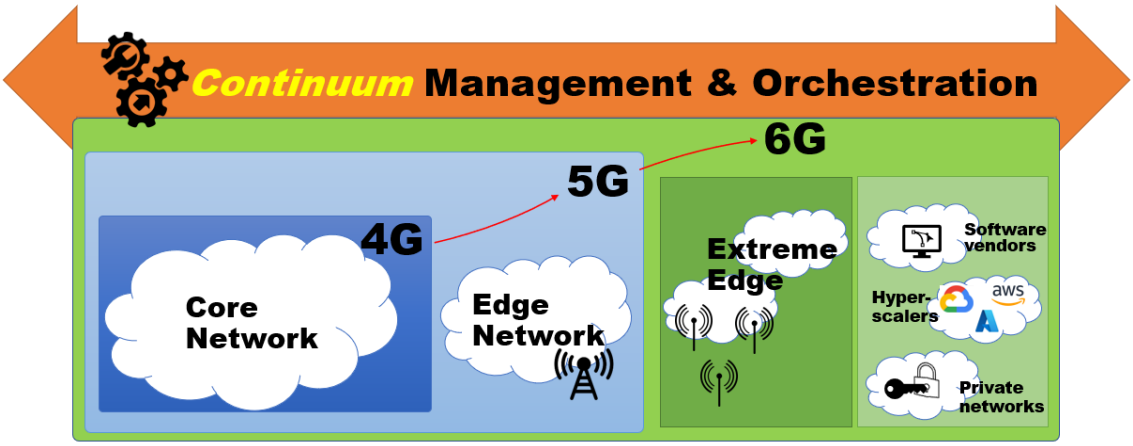

Figure 3-11: ***Continuum Management and Orchestration***

As shown in Figure 3-11, M&O in the 6G mobile networks considers not only those traditional network domains under the direct control of the MNO, as already considered in previous generations (core network in 4G and before, and edge network in 5G), but also the new network domains managed by

other entities and actors. Examples are the extreme edge (the end-user's domain beyond the access network)[6] and other 3rd party networks (e.g., hyperscaler networks, private networks, Software (SW) vendor networks, and others).

While it is true that in 5G there can already be certain network services that could communicate with the end-user's domain, or with other external network domains, the objective in Hexa-X is more ambitious. Hexa-X considers continuum M&O as a built-in architectural feature, integrating these new domains as an additional set of common infrastructure resources on which network services can be orchestrated.

One of the main challenges derived from this innovation is the integration of the extreme edge domain. This domain has very different characteristics compared to those datacentres deployed in strictly controlled environments on which 4G or 5G services are commonly deployed. The extreme edge refers to the end users domain, where there can be a huge number and diversity of devices (personal devices, industrial devices, entertainment devices, IoT devices…), with a diversity of supporting technologies, some with quite limited computing, storage, and power resources, others with highly dynamic behaviour. In summary, we can consider this for some parts a quite asynchronous and heterogeneous environment, which obviously poses new challenges regarding the infrastructure inventory synchronisation (that should be performed in a very dynamic way) and the deployment of the network services themselves. The continuum M&O, however, is not necessarily monolithic and may be composed of multiple, federated M&O systems of different technological or administrative domains.

Besides the integration of the end user domain, the optimal placement of NFs in such an extensive and diverse infrastructure is also a challenge by itself (More details can be found in Section 3.3.4). This functionality is in fact a well-known NP-hard problem [SSS11]. To address this, it will be necessary to design and evaluate efficient mechanisms focusing on optimising the dynamic NFs placement and the infrastructure resources optimisation.

Another obvious challenge regarding the M&O integration of different administrative domains is the definition of the necessary external interfaces to enable the interaction with other external orchestration platforms/frameworks that are out of the MNO scope and are managed by their own administrative entities. Along with the MNO network itself, it must be possible to deploy and orchestrate the network services over the different networks in a regular, reliable and secure manner (this would be also in line with the "network of networks" concept, which is one of the main abstractions in this Hexa-X project, see also Section 3.3.2 [HEX21-D51).

To make all this work together, new service description models should be provided in order to ease the definition of heterogeneous multi-domain network services, as well as the associated management tasks, considering the extended heterogeneity and complexity of the infrastructure. These new service description models should enable the orchestration of a wide variety of service definitions and decompositions, including physical elements and legacy virtual appliances, but mainly container-based microservices following the latest cloud-native architecture patterns, and serverless functions in all domains.

The main enabler to realise the above-described vision is automation, which can help to reduce configuration errors, handle complexity, reduce the time to create new services, and perform continuous resource usage optimisation, among others. Further, this should be done in all the network domains, first in each single network domain and then from an E2E perspective across the Device-Edge-Cloud continuum. This is relevant if we consider that in such a multi-domain context it will probably be necessary to handle a massive number of NFs, which could go beyond the human scale in terms of management and orchestration. Control loops should be applied to the whole lifecycle management of a Network Service (NS), from instantiation to decommissioning, including scaling, update, online orchestration actions (e.g., NFs scaling, placement, or configuration) on the services already running

[1] This is also sometimes referred to as "Far-Edge", "Deep-Edge" or "Fog Domain".

with minor or even no human supervision (zero-touch), as well as to perform Continuous Integration, Delivery and Deployment (CI/CD/CD) of the developed software, based on the cloud native principles.

This increasing degree of automation should be accompanied by a high level of network programmability. The system must provide programmatic interfaces on all the network segments in order to facilitate the implementation of the automatic M&O procedures.

Also, these network programmability procedures (or, at least, a significant part of them) should be accessible for all the involved stakeholders, regardless of their degree of knowledge of the low-level resources available in the network: the MNO, of course, but also other less specialised stakeholders participating in the ecosystem, such as verticals, content service providers, or even end-users. The degree of accessibility may vary depending on the stakeholder, especially depending on the nature of the stakeholder and also his willingness to have access to such freedom (stakeholders with limited to no knowledge may not be willing to have too many responsibilities). In order to ease access to this type of users, intent-based mechanisms should be provided for describing requirements, diagnosing the performance, modelling/abstracting the services and networks, as well as implementing corrective actions, among others, and all that using high-level declarative abstractions.

Another key enabler for network automation and simplified troubleshooting is a flexible monitoring and diagnostics system, able to provide information regarding QoE/QoS and infrastructure metrics. This system should be able to collect and distribute data from/to the different domains. Also, in order to ease the programming of useful and efficient automation control loops, this monitoring and diagnostics system should be well integrated with the automation system. A relevant feature associated with the monitoring system would be the possibility to provide mixed application and infrastructure-based metrics, in order to ease self-adaptation and self-optimisation decisions. We consider this feature as an innovation with respect to the state-of-the-art orchestration systems, which typically focus only on infrastructure metrics. Moreover, the monitoring system should be enhanced by analytics engines able to detect anomalies and predict future values of KPIs.

3.6.2 AI-driven orchestration

The previous section has introduced a number of orchestration challenges to reach the full vision of the Continuum Orchestration. One of the key enablers to efficiently address some of these challenges is the pervasive adoption of data-driven and closed-loop approaches towards higher levels of network automation. In this context, the AI/ML techniques are appearing as a promising solution to support the M&O system complexity mainly in different areas, e.g.:

- Time Series Processing. This is one of the common applications of AI/ML techniques [LIM21]. In our case, infrastructure and application metrics provided by the monitoring system can be seen as a complex set of heterogeneous time series. These time series would be processed by AI/ML algorithms in different ways, e.g.:
  - By correlating them using unsupervised learning algorithms to find hidden patterns in users and network behaviours. Then, the outcome from these correlations could be used to trigger M&O actions (e.g., NFs scaling or placement actions).
  - By making predictions based on the detection of regular users or network behaviours. This would be applied for implementing proactive orchestration strategies (e.g., in-advance network slices dimensioning by predicting demand and/or resource utilisation).

- Regarding the integration of different network domains, AI/ML may be used to address the following issues:
  - To support the integration of the extreme edge domain. As mentioned, this domain represents a quite diverse and potentially huge ecosystem with many different device types, supporting technologies, and information models. AI/ML can be a valuable asset in this context also in different ways, e.g.:

    - For performing the integration and normalisation of the heterogeneous data from this extreme edge context using big data analytics [LLF+21].
    - By applying federated learning techniques [YYY+19], which can be used to delegate certain learning processes on the end-user's equipment.
    - AI-Agents [AAFO] could be distributed on that infrastructure to gather metrics and/or perform local orchestration actions.
  - To support the complexity associated with the NFs placement problem. This is related to the optimal usage of the available infrastructure resources and, as mentioned before, is a well-known NP-hard problem. Although there are already certain state-of-the-art algorithms for addressing this [GKM18] [AMB20], AI/ML techniques have also demonstrated good performance on resolving this kind of optimisation problem [TCP91] [JAS02], so they could be a better choice for addressing this (we have to consider that in our case the infrastructure resources set can be quite dynamic due the integration of the extreme edge resources).

- Support the Operations Management processes. In this regard, there is a specific methodology known as AIOps [DLH19] focusing on using AI/ML techniques to simplify operations management and accelerate/automate problem resolution in complex modern ecosystems. This methodology can be applied in different ways, e.g.:
  - To gather and aggregate the huge volumes of operational data generated from the multiple infrastructure components, network services, and performance-monitoring tools. Considering that in a cloud-native environment the number of deployed NFs can be huge, AI/ML can provide valuable mechanisms to better handle the scaling of data collection and pre-processing.
  - For intelligent alarms filtering. In complex ecosystems like this, operational teams could receive multiple simultaneous alarms on their consoles in case of an incident. This could make it difficult to perform the incident analysis and determine the root causes. AI/ML has proven to be useful also in this regard [CRA+21].

- Support the intent-based configuration mechanisms mentioned above[7]. As a whole, intent-based techniques consist of defining the network state in a declarative way (stating the desired end result) rather than in an imperative way (stating very explicitly each single configuration action required to achieve a goal). This implies enabling the possibility of performing complex configuration actions using a high-level language, even close to natural language. As it is well known, AI/ML techniques are commonly used to perform this kind of task [SZF+18] [SZI21]. Moreover, AI/ML techniques can be applied to automatically recognise an evolution of the service intent and/or to verify the correct matching between the applied network configuration and the initial intent, dynamically updating the intent translation criteria when required.

3.6.3 Alignment with the Hexa-X E2E architecture

As mentioned at the beginning of this section, the core work related to M&O is being addressed in a separate work package (WP6), where the architectural design of all these novel orchestration and management mechanisms is being carried out aligned with the milestone MS5 of the project. However, both architectural designs (i.e., the general E2E architectural design and the specific M&O architectural design) are likewise based on the cloud-native design principles, so there is a good alignment with each other.

[7] Intent-based mechanisms is one of the main tasks in WP6.

4 Security, Privacy, and Trust

D1.2 [HEX21-D12] introduced two key aspects for the development of the 6G architectural proposal of Hexa-X: *Level of Trust* (LoT) as the KVI indicator related to security and privacy, and a set of *enablers* supporting these security and privacy goals. As part of the architecture evolution described in this document, this section goes into more detail on the LoT concept and its evaluation, and on how these security enablers fit in the proposed architecture to achieve the guidelines provided by the security considerations in D1.2.

While the focus on the following discussion will be on the 6G-specific aspects regarding security and privacy, it will consider the impact of next-generation networking on the application of fundamental security techniques as well.

4.1 Security Architectural Components

Figure 3-1 shows the overall architecture, visualising the applicable security and privacy components in all areas, and highlighting the specific 6G security technology enablers introduced in [HEX21-D12], further discussed in this section. While the focus lies on the technology enablers as new architectural components, a holistic 6G security architecture must also comprise today's well-proven security mechanisms, as far as they are still relevant in 6G. We briefly summarise them in the following figure, without aspiration of exhaustiveness and depth of detail.

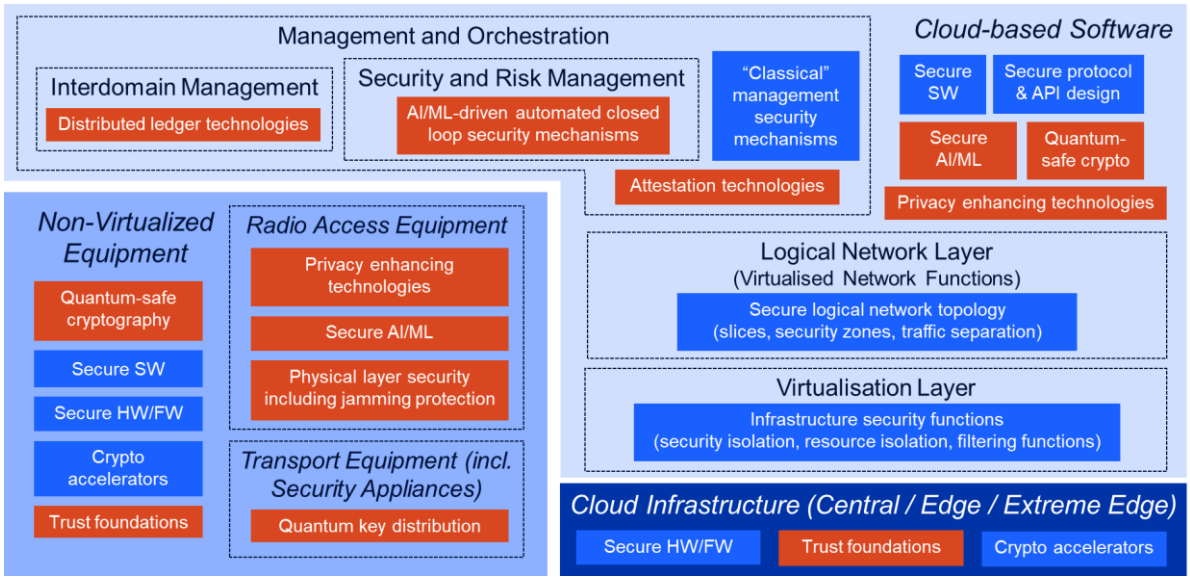

Figure 4-1: ***Overview of the essential 6G security architectural components***

Figure 3-1 distinguishes among non-virtualised equipment (for radio access and transport), the cloud infrastructure, and the software running on it, including the virtualisation layer, the logical network layer, and the management and orchestration functions, including security and risk management and interdomain management. In each part, the figure shows the most relevant security and privacy building blocks (or architectural components), with the new 6G security technology enablers highlighted in red, and the more traditional building blocks like for example "Secure SW" in blue.

Many building blocks apply to multiple areas, e.g., "Secure SW" applies to the non-virtualised radio and transport equipment (as far as this equipment comprises software), to the virtualisation layer and all the software running on it, including management and orchestration functions. As another example, "Trust foundations" apply to all hardware, i.e., the radio and transport equipment as well as the cloud infrastructure. On the other hand, some building blocks appear at dedicated places only, like "Distributed ledger technologies" appearing at interdomain management only, but this does not

preclude the potential applicability of the building block in other areas. Also, when a building block appears in an area, this does not imply that the building block is always applicable. For example, certain non-virtualised radio access equipment may not have access to sensitive data, so no privacy enhancing technologies may be required here. As another example, obviously not all transport equipment is required to support quantum key distribution.

The traditional security building blocks may be mostly self-explaining, but note the following:

- "Secure SW" refers to software with a low (close to zero) degree of vulnerability. "Secure HW/FW" has the same meaning for hardware or firmware. An example is the robustness of a processor against leaking information between different processes running on this processor in a (quasi-) parallel manner.
- "Secure protocol and API design" refers to robustness not only against external attackers (which is typically achieved by the use of cryptography), but also against erroneous or malicious behaviour of authorised peers.
- "Classical" management security mechanisms" comprise well established mechanisms such as access control, role-based access, secure logging, isolation of management functions/traffic from all other traffic, etc.

4.1.1 Trust Foundations

Trust foundations are hardware-based tools used for the root of trust establishment. These tools are listed as (i) trust anchors comprising Trusted Platform Modules (TPMs) (for software integrity) [ISO15], trusted execution environments (TEEs) (for isolation against the infrastructure provider) [CCC20], hardware security modules (HSMs) [HSM], (ii) secure identities, crucial to build trust between UE and network and (iii) secure attestation methodologies [ERI20]. They can be employed at different architectural components.

Trust anchors can exist in radio access components (e.g., secure environments in base stations according to 3GPP TS 33.501 [33.501], to provide protection against physical tampering). Trust anchors can provide protection against external adversaries for virtualised radio functions, e.g., central unit functions when RAN functions are split into Distributed Units (DUs) and Central Unit (CU).

With 6G, it is envisioned that edge and extreme edge computing shift part of the computation and intelligence from the core network to heterogeneous edge devices (smartphones, wearables, sensor networks, connected cars, industrial devices, connected home appliances, etc.). This shift may also pave the way for the advancement of trust foundation technologies. Although secure identities, e.g., SIM, embedded SIM (eSIM), or integrated SIM (iSIM), are existing options to be used as the root of trust to meet access control requirements, there will be new challenges in edge intelligence, incorporating ML models trained and deployed on the edge and extreme edge devices. For example, model integrity, model robustness, model protection, and training data privacy need to be addressed in these extreme edge devices that can be resource constrained IoT devices.

At the transport layer, the fundamental trust anchor that can be adopted in switches can be TPMs to protect software integrity. HSMs and secure identities can be used to establish (automated) secure communication and access control mechanisms.

To address the virtualisation layer and logical network layer requirements, like confidentiality of sensitive data on certain network functions which should not be reached by other network functions, TEEs and HSMs can be used to provide appropriate protection. Secure attestation of virtual network functions is needed to verify isolated execution and to ensure that a network function behaves as expected based on the specification of the function.

Trust anchors and attestation techniques become very important in the usage of the cloud for the deployment of entities and functionalities to ensure that the sensitive data is not leaked, there is no intervention to the process and the running software is the right one. Trust anchors and attestation

techniques can be utilised to separate management and orchestration from the other communication between network entities.

4.1.2 Privacy Enhancing Technologies

Privacy Enhancing Technologies (PETs) is an umbrella term for advanced building blocks (but not limited to), namely differential privacy [DR14], homomorphic encryption [RAD78], secure multi-party computation [YAO82], and Confidential Computing (CC). PETs are used to bring advanced solutions to the challenges that cannot solely be handled by classical privacy techniques such as anonymization, or consent mechanisms.

6G will allow high precision and accurate localisation/positioning and sensing that enables new emerging location-based services (see also Section 3.2.3). This also opens new privacy challenges from the UE perspective e.g., tracking and locating UEs would be more precise and with more detailed data and context. New accurate sensing capabilities may lead to misuse of the sensing information when it is used beyond the granted context and violates UE privacy, thus may cause unwanted effects, e.g., environment/scene may include more info than intended or this info might be used maliciously. Privacy by design approach [GDPR] should be followed during the design of localisation/positioning and sensing functionalities in 6G [HEX21-D31]. Privacy enhancing technologies, as well as data protection techniques enforced by regulations, need to be adopted.

Some of the networks functions store credentials used for the establishment of a root of trust. For example, functions similar to the current Unified Data Management/Authentication Credential Repository and Processing Function (UDM/ARPF) hold the Home Network Public Key Identifier(s) for the private/public key pair(s) used for subscriber privacy [33.501]. The protection of these keys becomes a vital issue because if it is compromised then the prevention mechanisms to protect subscriber privacy will not achieve their goal. Another example would be some other functionalities similar to UDM/ARPF and Unified Data Repository (UDR) network functions which store and/or process subscription credentials such as the long-term keys used for authentication and key derivation to secure the subscriber communication to the 3GPP network both for control plane and user plane [23.501]. Thus, it is very important to protect the credentials that are considered as the root of trust for the security and privacy. In cloud deployments, these credentials may not be wanted to be revealed to the cloud providers in cleartext. For this point, confidential computing, secure two-party protocols, and homomorphic encryption solutions could help to remove privacy and zero trust-related barriers to overcome the problem of allowing the use of cloud technologies.

For the management and orchestration, user (UE) data may be utilised to improve network management. To be able to use user data without any violation of user privacy, privacy enhancing technologies can serve as a key enabler. Better security and risk management, cooperation among different operators and actors in the field for threat intelligence, fraud detection, or other possible applications could help security analysis and predict future risks. However, there could be some concerns about sharing data among actors. To resolve these concerns, privacy enhancing technologies could be used in collaborative AI/ML methods (e.g., federated learning, split learning) [HUA21], [KOB21].

4.1.3 AI/ML Assurance and Defence

In the design of the 6G air interface, AI/ML enablers are proposed for intelligent, more capable, sustainable radio access networks. AI-based air interface design pillar, requirements, and gap analysis are comprehensively studied in D4.1, [HEX21-D41]. AI/ML enablers are foreseen in data-driven transceiver design approaches for hardware impairments [HEX21-D41], AI-driven transmitters for beamforming optimisation, and fast initial access. Receiver side functionalities like channel estimation can leverage AI/ML. Since radio interface functionalities are changing with the cell-free and distributed MIMO systems [IBQ+19], new areas to embrace AI/ML will emerge. The learning approach in 6G is heading towards more distributed processes. Especially the combined design of the communication and computation with the aim of benefiting from the multiplicity of nodes available will allow for using distributed resources in the network. Collaborative learning methodologies such as federated learning will not only pave the way for the use of distributed resources but also improve privacy by addressing

the data minimization aspects, as data will not need to be transferred from distributed resources to a centralized environment.

Solutions to protect the robustness and privacy in non-collaborative and collaborative AI/ML are inevitable to achieve security assurance in AI-driven air interfaces. Attacks performed by distributed and sometimes maliciously behaving network elements and also by UEs should be avoided. For example, when UE location information is used in a collaborative setting to predict network parameters, a targeted membership inference attack may reveal the information whether a specific UE is present in a certain location, thus, causing a privacy violation. Another implication can stem from a UE sensing application when malicious UEs poison the training input or disrupt the model behaviour via adversarial examples to reduce the model robustness [GSS14]. Threats associated with the training and inference pipelines including data poisoning, model inference attacks, and model reconstruction [TBH+19] should be carefully identified and addressed.

Moving to transport networks, AI/ML mechanisms can be used to automate maintenance, management, and monitoring purposes on the transport layer. Edge intelligence capabilities enable improved management and scheduling of edge resources using AI/ML to provide efficient task and service processing. AI/ML applications can be deployed to provide AI-as-a-Service (AIaaS) on central cloud or edge. In these cases, protecting model parameters and inference APIs in a virtualised/cloud native environment becomes essential. Model extraction or model inversion attacks might be possible using query results received through the inference APIs. To protect against these attacks, differentially private model disclosure and output sanitisation techniques can be employed [JKM+21]. Confidential computing protects against model stealing attacks on AIaaS deployments. Well-known security control mechanisms such as communication security, access control, monitoring, and logging also provide mitigation against some attacks performed over inference APIs. To provide security assurance for consumers of an AI service, attestation technologies can be used.

It is foreseen in the future that AI/ML plays a vital role in security management and orchestration with respect to different elements of networks such as cloud infrastructure, microservices, virtual/physical network functions, and network slicing. To address this, the ETSI ISG ZSM proposed Zero-touch [ETS19] network and Service Management (ZSM) architecture is expected to further evolve to achieve network automation in 6G networks. The ZSM architecture is formed by modular characteristics using intent-based interfaces, closed-loop operation, and AI/ML techniques to empower full automation of the management operations. Intelligence in network operations can be achieved by AI/ML on two levels. Firstly, by the network itself, through AI-enabled self-configuration/self-optimisation/self-healing operations, leading to a steady self-organising network. Secondly, by the verticals or developers deploying their services and applications on the network, with Development, Security, and Operations (DevSecOps) powered by AI/ML. Therefore, it is important to develop robust and resilient AI methods to be used in network management and orchestration as well as in security analytics for fault diagnosis and providing recommendations throughout the development lifecycle.

Furthermore, AI/ML-driven security enablers can be deployed in security and risk management for radio, transport, and data centre equipment including security functions in the virtualisation and logical network layer, and in the management layer itself. When a network gets attacked by the injection of malicious traffic, it can affect the E2E service management domain as well as the management sub-domains. The detection of such attacks in an automated and distributed manner and executing the mitigation/prevention plan are complex processes in a multi-domain network service management architecture (e.g., ZSM [ETS21b]). Cognitive anomaly detection, attack mitigation and prevention mechanisms based on AI/ML can be incorporated for automating the security operations in 6G architecture with the automated closed-loop security management operations such as monitoring, threat intelligence, analysing, and intelligent decision making.

4.1.4 Quantum Security

Different scenarios are possible related to the success and adoption of the quantum computing and communications technologies that will impact 6G networks and therefore should be considered as a

source of the attack, but also as a solution. The following list covers the main trends in this security area:

Quantum-safe crypto. D1.2 introduced how quantum computing capacity [HPE+] will reduce the effectiveness of different cryptographic 5G algorithms. Quantum-resistant cryptography algorithms to replace current algorithms are in the study and will be considered for 5G-Advanced and 6G. Nonetheless, the approach should be progressive. For example, the explicit use of keys of 256 bits of length for the key derivation process, cyphering, and integrity/MAC could be enough in a post-quantum era for 5G-Advanced. Also, one advantage in 3GPP standards is the possibility to migrate to Post Quantum Cryptography (PQC) algorithms keeping backward compatibility. Terminal to USIM communication is based on a common interface since 3G (to provide cipher key Ck and integrity key Ik) and probably also for 6G, so the use of PQC should not change the interfaces, requiring only to update ARPF and USIM (or eSIM) internal code to support new algorithms to provide the key derivation functions in the AKA specification (functions f1 to f5) [35.231] [35.206]. Asymmetric cryptography requires more attention. For example, the Subscription Concealed Identifier (SUCI) uses the Elliptic Curve Integrated Encryption Scheme (ECIES) algorithm, which will make it easier to obtain the private key from the USIM or operator issuer, reverting to 4G security weakens in exposing user identities. To avoid privacy and user tracking problems (e.g., IMSI-catchers [PGB+21]), PQC algorithms should be adopted in the 3GPP standards to appropriately protect SUCI in 6G. Also, network domain security (backhaul transport, roaming) is relevant and impacted by quantum threats. Cypher suites in Layer 2, 3, and 4 security protocols (e.g., MACsec, IPsec, TLS) should be appended with PQC and combined with alternative quantum-safe key agreements mechanisms (e.g., Quantum Key Distribution (QKD)), or provide hardware security and avoid future unknown attacks in PQC (e.g., new quantum algorithms or vulnerabilities in the implementation).

Quantum Key Distribution (QKD) is considered a quantum-safe mechanism for key agreement and key generation (see [HEX21-D12]) in optical links. Some relevant examples in 6G cover connectivity between future Radio Access Network (RAN) components (RRU, DU, and CU) and distributed core (edge, cloud, and operator datacentre) where quantum key distribution channels coexist with classical data and signalling channels over the same fibre [MAS+19], and where nodes are mutually authenticated using trust foundation technologies, such as TPM device authentication, or in combination with PQC. Current technology evolution offers standard interfaces to meet application demands, adopting advanced Key Management Systems (KMS) to provide customised key delivery, such as high key rates for one-time pad symmetric encryption or PQC algorithms at line speed. As a result, different 6G potential use cases could dynamically adapt their secure key demand, for example activating higher ciphering capacity between local trust zones, or ad hoc keys request needed to cover privacy information in human-centric communications. The natural integration of QKD systems and KMS within SDN architecture using standardised southbound interfaces over general optical and transport network controllers [ALL+19] will allow complex and dynamic deployments of quantum paths for key distributions initially over trusted nodes and, when they are ready, using quantum repeaters. Additional functionalities to be developed in parallel with 6G technology, such as point-to-multipoint key distribution, miniaturisation into optical pluggable, trusted nodes/quantum repeaters, and multi-vendor compatibility, will allow covering key distribution and rekeying policies for security high demanding 6G internal services (e.g., lawful interception, secure time synchronisation) and external services, e.g., Edge applications, sensible Industrial IoT devices, AR/VR.

Quantum Random Number Generators (QRNG). Cryptographic algorithms, including classical and PQC, depend on high entropy randomness to guarantee security. The technology to generate true random numbers based on quantum principles is evolving to miniaturised solutions [IQQ] to be integrated into the hardware and delivered with solutions that would be used in mobile authentication solutions in 6G, such as RAND value in AKA. Also, alternatives such as Entropy as a Service (EaaS) [VAS16] could provide the needed randomness in case of massive IoT adoption related to the extreme edge where weak internal RNG could compromise the security.

Quantum sensing and metrology. Improvements in these technologies could lead to robust solutions increasing the resilience of time synchronisation in future 6G networks, using fibre, as a complement to GNSS actual system [GSM21].

4.1.5 Distributed Ledger Technologies

Distributed Ledger Technologies (DLT), especially Blockchain, have been actively explored on how to ensure security in service providers or telco operators internally, and externally in multi-domain interactions [AGB+20, FBF+]. Particularly this may include the secure intra/interdomain management procedures in next generation networks in the management and orchestration planes. The use of smart contracts enables broader options on where to apply Blockchain for ensuring security.

A useful example is the creation of dynamic Service Level Agreements (SLAs)/Secure Service Level Agreements (SSLAs) for E2E multi-domain network slices. Service providers are able to offer verticals on-demand rapid deployment of network slices that provide user access in (geographical) areas where the service provider has no service footprint (e.g., different country, continent) [UZK+]. With the use of Blockchain and smart contracts, service providers are able to securely discover and establish E2E slices that would span over multiple administrative domains in a matter of seconds, thanks to the lack of a centralized entity nor the need to reach an agreement a priori among the participating domains.

The inter-domain interaction can go through public (permissionless) Blockchain, private (permissioned) Blockchain, or both. Permissionless Blockchain allows a more open provision of Distributed Ledger services, with longer computation requirements that may downgrade the efficiency of the interaction. In the other case, permissioned Blockchain realisations limit the provision of ledger services to participants in a closed community that share trust links, implying a lower computational overhead that may boost the performance for applications. Using both realisations, service providers are able to establish layered interactions between different entities and domains. For example, permissionless DLT/Blockchain can be a trust-enabler for applications that require secure interaction in dynamic solutions where different entities or customers are unknown to each other. Or DLT-based secure common marketplaces (e.g., brokers) can be developed to use the so-called *crypto-currencies* for the sharing of virtual and physical resources [BGL+]. In parallel, a permissioned DLT/Blockchain can serve to provide a secure platform for a consortium or a group of well-known administrative entities for a secure interaction in establishing E2E network slices, sharing infrastructural resources, or sharing data.

An important advantage of applying DLT is the increase of distributed services which avoids single points of failure like in centralised platforms. Each service provider may participate in multiple ledgers, using multiple nodes. Multiple ledgers enable the layered interaction, while multiple nodes increase the platform security and resilience of the network (e.g., Distributed Denial of Service (DDoS) avoidance) [RBL+17].

The append-only nature of DLT enables high transparency and accountability of the participants [ETS20]. Therefore, any interaction among service providers, or among vertical customers via a DLT solution contains a transparent history of records without exposing the private data while increasing the trust and accountability.

4.1.6 Physical Layer Security

Physical Layer Security (PLS), as discussed in D1.2 [HEX21-D12], can be applied as a new security measure in addition to current mechanisms, which generally are of cryptographic nature. The use of PLS for secret key generation, for authentication, and for radio link integrity monitoring is discussed in the context of wireless communication and thus might be an option for 6G as well. In particular, the latter option allows synergies with joint communication and sensing concepts as they are discussed for 6G. Since PLS does not rely on computational complexity, it would provide quantum-resistance with low costs of time and energy. These outstanding features would make PLS a competitive solution that should be identified as a candidate component of the 6G security architecture. Additionally, the nature of PLS, i.e., that it purely focuses on the physical properties of the channel, would also release it from any dependency on other security mechanisms, such as key generation and distribution, or any conflict

with them. However, it must be considered that a deeper understanding of the implications of exploiting the use of PLS in 6G is required, and this is currently beyond the scope of the project.

4.2 Level of Trust

Deliverable D1.2 [HEX21-D12] introduced the concept of Level of Trust (LoT) as the KVI to be considered to assess the security of a network service in a particular application environment. Such a LoT, as much as it happens to the QoE (Quality of Experience) concept [MUS17], cannot be directly measured or calculated by a general formula from common measurements, and has a subjective component, but necessarily it is:

- Associated to the *specific context* in which the network service is being consumed.
- Evaluated according to the *user intent* when requesting the service.
- Derived from evidence collected during the *service lifecycle*.
- Suitable to be assigned or audited by *independent third parties*.

The evaluation of the LoT applicable to a requested service can therefore be associated with two different stages. In each of these stages, the necessary matching among policies and requirements, analysis of evidence, and interaction with users, including the subjective component mentioned above, would require the use of AI capabilities. In addition, evidence and formal agreements, whenever required, would imply the use of reliable decentralised mechanisms for data sharing and access, typically by means of DLT approaches. In the first stage, an *assessment* of the achievable LoT is made, according to the combination of the features required by the user and the available policies and measurements on the infrastructures and network functions to be allocated for service provisioning. This stage corresponds to the phases for service deployment, selection, and activation, and implies the availability of negotiation mechanisms for network service consumption (like network APIs) or the availability of a formal security policy for predefined services.

In the second stage, an *evaluation* of the actually achieved LoT is performed, as a result of the evidence data collected, the key indicators selected in the assessment stage, and whenever applicable, the enforcement of the automated mechanisms established by the service level agreements settled in the previous stage. This stage corresponds to the phases associated with service consumption, assurance, and finalisation, and assumes the availability of data infrastructures able to collect and aggregate relevant monitoring data for further auditability.

In each of these stages, the necessary matching among policies, and requirements, analysis of evidence and interaction with users, including the subjective component mentioned above, would require the use of AI capabilities. In addition, evidence and formal agreements, whenever required, would imply the use of reliable decentralised mechanisms for data sharing and access, typically by means of DLT approaches.

Work on trust assessment and trust models for next-generation network infrastructures is an active research field [VSP+22], and the Hexa-X security team plans to leverage recent results to achieve a better understanding of trust evaluation mechanisms. As an initial approach to this task, a concrete procedure for LoT evaluation will have to identify:

- Actors in the telecom infrastructure: Who is/are the trustor/s (the one that is trusting) and who is/are the trustee/s (the one that is trusted).
- Baseline requirements: fundamental technologies/controls that are in place.
- Legal requirements as enforced by the legal authorities.
- The availability of advanced security technologies, e.g., the existence of TEEs, use of attestation (The possible use of trustworthy hardware and/or software for the creation of attestations could be investigated).
- The need for trusted third parties (e.g., in electronic transactions TTPs being service providers can provide validation services).

4.2.1 Achievable LoT Assessment

In the first stage, we can identify the following sources for evaluating the achievable LoT:

- Policy declarations for services.
- Infrastructure capabilities: versions, features such as roots of trust or secure execution, support for attestation, attestation statements, etc.
- Function capabilities: versions, features, attestation statements, etc.
- Monitoring and verification capabilities: slice isolation measures, topologies, etc.
- User intent, in a controlled vocabulary, as the selection and parameterisation of policy declarations and capabilities.
- Records of previous performance: incident records, user experiences, etc.
- Records of previous performance for (somehow) trusted third parties, in what constitutes reputation statements.

The expected result will be an assessment on the achievable LoT, including:

- The selection of a particular offering, satisfying the expressed intent.
- The identification of relevant metrics to evaluate the LoT in the evaluation phase.
- Optionally, a set of smart contracts derived from matching declaration, capabilities, and intent.

4.2.2 Achieved LoT Evaluation

In the second stage, these are the sources to be used to evaluate the achieved LoT:

- Available infrastructure and functional features, according to actual resource reservations and log entries.
- Relevant service and infrastructure telemetry data, as registered by trustworthy data repositories.
- Results of specific measurements executed on-demand.
- Recorded events: incidents, alarms, etc.
- Optionally, smart contract triggering.

The expected result will be a record of the achieved LoT, including:

- Aggregated data for further auditability and reputation assessment, including privacy protection.
- Feedback about intent statements and matching procedures, including possible alarms related to mismatches, as a backtrack of the experienced events and collected metrics.
- Optionally, execution of the agreed smart contracts.

4.3 Security Management

The services promised by 6G are of little benefit to customers unless they are defended against cyber threats with appropriate safeguard measures. To deploy an effective and consistent global defence with limited intrusion to avoid the degradation of the end service performance, it is required to consider the security services at all stages of the 6G network's management life cycle, from the preparation stage to the decommissioning of the network [5GP21]. The security management has many similarities with the regular management of network functions, but also specificities related to the nature of this management. It aims to protect the valuable assets of the network such as the continuity of customer services, preserving the confidentiality, integrity, and availability of the supporting assets such as data, functions, and resources, while being compliant with cybersecurity regulations and industry standards. To achieve this goal, the security management has to be designed in accordance with the principles defined in the public or government cybersecurity frameworks. Examples of frameworks include the

Security Incident Management Guide for Computer Security Incident Response Teams [RWT13], the Managing cybersecurity for Industrial Control Systems guide [ANS21], or the NIST cybersecurity framework [NIS18]. Although each framework has its specificities, they all follow a very similar pattern.

In this section, we discuss how the security management can be integrated within the general management and orchestration architecture of 6G, how it can be automated, and what are the challenges to overcome.

4.3.1 Integration within the general management

6G is expected to raise the heterogeneity of the network to an unprecedented level [ZSV+21]. To manage this complexity while maintaining flexibility, the management should be organized in a distributed and hierarchical manner. Lower-level management entities would manage a dedicated, criterion-driven subset of the entire system to more accurately analyse data and react more promptly, while higher-level entities would perform E2E management to properly deliver the requested service [5GP21]. The security management should follow the same hierarchical organization, for the same reasons. In practice, this means that any set of managed entities should have an associated security management entity, which is in charge of controlling the security processes. A security manager can monitor the health status and receive an activity report from both managed objects and other managers, and trigger actions back to them.

An example of such integration of a security manager to a general manager is proposed by ETSI ISG NFV in the context of network services [ETS21]. However, this manager is not integrated within a hierarchy for E2E capabilities, as ETSI ISG NFV focuses exclusively on the network service level.

4.3.2 Automation

Similarly, to generic management, security management should have a high degree of automation in order to manage the complexity of 6G systems in a dynamic and flexible way. This automation is often envisaged as being provided by AI/ML algorithms embedded in closed loops [5GP21] [ZSV+21] which also constitute the core of security managers. To further increase the flexibility of the traditional closed loops, such as MAPE-K [IBM05], it is possible to integrate them into a service-oriented framework with hierarchical capabilities. For example, ETSI ISG ZSM [ETS19a] proposes such a framework which supports service-oriented, multi-domain, and multi-tenant features. Considering the security aspects of ZSM, ETSI ISG ZSM is in process of conducting a security risk analysis of the system [ETS21b], although no detailed implementation proposal has been produced yet.

4.3.3 Challenges

The heterogeneous and dynamic nature of the 6G network, involving multiple stakeholders as well as strong automation requirements, poses a number of challenges for the security management system.

The first challenge is the massive amount of monitoring data collected across the widely heterogeneous and distributed 6G system to be analysed to verify the expected level of security. Security enforcement requires to partially or entirely analyse the traffic that crosses the network, and this process can consume a lot of resources. To mitigate this problem, monitoring can be dynamically adjusted to the context, both in terms of frequency and the number of targets to be monitored [Hue08]. This would effectively create a nested control loop dedicated to intelligent monitoring, within the general control loop. Such autonomous monitoring must be carefully designed to resolve the trade-off between available resources and the level of security required.

Another challenge lies in automating the first two core functionalities of the security framework: Identify and Protect. These two steps consist in making an inventory of valuable and supporting assets, analysing threats, vulnerabilities and risks, choosing risk reduction strategies, and deploying all appropriate security measures (dedicated functions, configurations, network topology changes, etc.). Compared to the other framework functionalities - Detect, Respond, and Recover - there is only limited

research on automating the Identify and Protect functionalities which pose a significant challenge to the goal of fully automating the framework processes.

Automating security management implies having a common security language between the security manager, its clients, the underlying resource manager, and the available tools to perform security actions, such as the Virtual Security Functions (VSFs). Between the security manager and its clients, the service request must allow the security manager to clearly deduce the security needs, even if the request is formalized as an intent. While clients' needs and service capabilities are usually well developed for the business services, for which clients have a clear idea of their needs and a strong incentive to express them clearly, security may be overlooked as it is not part of the business itself. Clients may have less knowledge about it, leading to either unclear or complete lack of description of security needs. Similarly, tools that are put at the manager's disposal, such as Virtual Security Functions (VSFs, must clearly detail the security measures they are able to provide, so that the manager can use them to fulfil the security requirements. This includes, for example, many security enablers that are detailed in this document, such as quantum-based security, as well as legacy security capabilities. Finally, when it comes to the instantiation of the service, the different resource providers must disclose information about the security guarantees that they can provide. This includes new providers envisioned in 6G, such as far edge, or even customers' premises, are envisioned to host services and functions. Those remote locations are unlikely to offer the same security guarantees as central clouds, as they cannot fully apply the security framework recommendations. This was already a concern for 5G, which was introducing the notion of edge. The owners of smaller infrastructures used by edge, far edge, or fog, which may include private small servers or even user end devices, do not have the same security capabilities as large datacentre owners. Typically, physical security, software updates, or hardware updates, may not be properly managed [RLM18]. These issues will have even more impact on 6G, which intends to rely on those kinds of infrastructures for key services, such as ultra-low latency services.

Regarding the security policies, the complexity of the 6G networks and the involvement of multiple stakeholders imply that those policies will be quite complex, and potentially issued by several entities. This raises the issue of managing the conflicts that are likely to occur between those policies, or between a policy and the current state or capabilities of the system, or even the inconsistencies that may exist within a single policy. This subject has recently attracted attention in the context of telecommunication networks and automated, virtualized environments [MBB+20] and has recently been tackled by the INSPIRE5G-plus European project [INS21]. However, this is still an ongoing research topic, and it should be investigated in the context of 6G networks.

As both functional and management architectures are envisioned to be service oriented, following for example the ETSI specified ZSM architecture [ETS19a], there will be an increased number of open APIs across the system. Furthermore, these APIs will be as dynamic as the services that implement them, as they will be created, deleted, and updated at the same time. As APIs are exposed to clients, they are common targets for attackers, so it is important to secure them. Although efficient security mechanisms already exist, such as TLS or OAuth2.0, they will not be sufficient to face the aforementioned API variety and dynamicity, as well as the volume of traffic passing through them [BT20a]. Since this volume can vary, it is necessary to adapt the size of the security functions to maintain a constant level of performance and thus avoid any disruption to the end service and waste of infrastructure resources. It may therefore be interesting to explore AI-based automation of the security of APIs, which would be carried out by the security manager. This automation would automatically apply adequate security measures to an API, such as DDoS mitigation or input validation, upon threat detection.

5 Spectrum evolution aspects

In this section, spectrum evolution aspects relevant to extending spectrum utilisation both in frequency ranges already in use (i.e., low, mid, and mmW) and in new frequency ranges (i.e., 100-300 MHz and above) to address 6G service requirements are addressed. Innovative concepts of flexible spectrum usage and management are presented as well.

5.1 Extending spectrum utilisation

To further optimise spectrum utilisation, technical enhancements are continuously being studied and specified for 3GPP technologies. Also, better spectrum sharing conditions of International Mobile Telecommunications (IMT) systems with systems of other Services could be achieved by taking more realistic technical characteristics and deployment scenarios into account.
In the following, a few examples of both spectrum utilisation and sharing enhancements are described:

- Improving the usage of available spectrum in the different IMT frequency bands can be achieved by enhancing spatial spectrum resources management through techniques such as advanced Carrier Aggregation (CA) and distributed cell deployments (i.e., cell-free/distributed MIMO) have been developed. By enabling devices to connect to a set of carriers available in network nodes simultaneously and flexibly, higher bandwidths can be achieved as well as an optimised spectrum usage across all available bands, e.g., spectrum reuse.
  For example, enhanced CA can enable scheduling data over multiple cells including intra-band cells and inter-band cells to ensure that FR1 scattered spectrum bands or FR2 (and some FR1 bands) wider bandwidth spectrum can be utilised in a more spectral/power efficient and flexible manner, thus, providing higher throughput and better coverage in the network. Additionally, in the distributed MIMO technology, a set of nodes act as one cell-less network enabling high-density deployment and spectral reuse. This is achieved by an efficient combination of antenna and transport solutions, which can more efficiently utilise spectrum resources through central coordination (this is also addressed in Subsection 3.4.1).
- New coordination mechanisms and techniques for local spectrum use would greatly improve spectrum utilisation efficiency. 5G local spectrum licenses have been introduced in certain countries to make spectrum available for different stakeholders to deploy their own non-public networks. Typically, local licenses awarding follows a “first come first serve” administrative allocation principle and imposes corrective actions in some form of separation distance for interference coordination with other networks, either local or nation-wide public. Maintaining the current interference coordination model would greatly impact the efficiency of spectrum usage, especially if local licensing would expand to hundreds or thousands of applying entities. In order not to deteriorate spectrum utilisation new mechanisms and techniques would be required based on feasible levels of information exchange between licensees to manage assignments and coordinate potential interference issues.
  Automated and standardised systems for efficiently getting temporary access to spectrum from national regulatory needs to fulfil many legal and regulatory requirements. Furthermore, technical enablers need to be defined to set up a fully automated temporary spectrum access lifecycle, including a spectrum request by a local spectrum user, spectrum granting by the national regulator, spectrum use and reporting by the local operator, and spectrum release for finalising the process. For example, in the desired geographic area and frequency band, the local spectrum user could perform a search for other users and transmit to the regulator classified results, e.g., by encryption, of this campaign in support to the request for spectrum usage in the

intended time and geographic extensions. The regulator would automatically send approval or rejection of the request, based on reserved centrally managed database information with a potential request for additional reports on periodic spectrum scanning reports. When this process is fully automated, lifecycle periods from minutes to hours are possible which allow to significantly increase the granularity and, with this, the usage efficiency of scarce spectrum resources. It is to be noted that such automated mechanisms are only suitable for specific use cases where Quality of Service (QoS) requirements allow for it. This is due to the fact that with these mechanisms, access to a fixed amount of spectrum bandwidth may not be possible at any time and without any delay. In addition, there are several challenges that need to be addressed such as the time it takes to detect the source of interference, the procedures associated with having a harm claim filed and reviewed, and the lack of observability of performance indicators associated with validating the interference claim.

- 6G Networks in Network (NiN) are prospective solutions that can allow interference-controlled operation in scenarios where one or several subnetworks are deployed in a larger, often wide-area network, using the same wide-area network spectrum. Subnetworks can be associated with vehicles, drones, machines, ships, trains, body area networks, or robots, but also be established by any coordinated moving group of sensors, e.g., a fleet of vehicles or drones. Subnetworks can be mobile, which implies frequent variations of distances between macro-network nodes (base stations) and subnetworks. The 6G NiN concepts comprise Intra-subnetwork (Intra-X), Inter-subnetwork (X2X), and subnetwork-to-wide-area network (X2I) connectivity. While X2I connectivity often relies on standard cellular links from a particular subnetwork node to the infrastructure, in the cases of Intra-X and X2X connectivity there is considerable work ahead if these are to be included in cellular specifications (e.g., they show similarities with D2D links but are not equivalent). In addition to technical and implementation issues (e.g., Intra-X radio resource management, Inter-X-aware interference-aware scheduling, etc.), certain NiN concepts may be relevant to the regulatory domain. For example, additional regulation, sometimes even safety relevant, for Intra-X connectivity (e.g., to completely shut off Intra-X links in case of interference)could be needed e.g., resulting in the requirement that radio resource management is performed internally, within the subnetworks as well, in a combination of a) Intra-X radio resource management, b) Inter-X interference-aware scheduling (based on AI), and c) maximum overall interference level, caused by an X to the wide-area network, controlled by the wide area network. In fact, interference issues between subnetworks and between subnetworks and wide-area networks will require careful coordination in order to avoid undue capacity losses.
- Improving incumbent receiver susceptibility to interference would be a key enabler for spectrum sharing and spectrum utilisation efficiency improvements. In identifying a new spectrum for IMT, coexistence between IMT and the other incumbent services needs to be assured and this is currently realised by means of measures mostly on the IMT transmitters side rather than, on a certain measure, on the transmitters side of the incumbent services. To this respect, for example in CEPT [CPT21], work is being done to scrutinise earth stations devices in use in the incumbent fixed satellite service at 3800-4200 MHz (e.g., microwave and intermediate frequency filters, Low Noise Block down-converters, etc.) to assess if a reduction of the separation distances may be applied while still protecting the incumbent Fixed Satellite Service (recognising that a number of satellite earth stations are currently using RF filters with higher selectivity and higher input compression points).
- Improving assumptions and models to better fit more realistic scenarios and not only pessimistic ones can lead to improved spatial separation requirements. Currently, separation requirements

are calculated assuming the worst-case conditions for both the new service nodes (BSs and/or UEs) positions and the propagation characteristics (i.e., the mobile service interferer is in the main lobe of the victim antenna and there are no obstacles that attenuate the interference). These conditions do not always hold; in fact, the incumbent service (e.g., fixed or satellite) victims typically uses antennas with very narrow beams, and not all BSs and/or UEs are in the worst-case situation. Consequently, smaller spatial separation and/or higher power could be used by IMT systems without causing excessive interference. Furthermore, it is useful to develop quality metrics for spectrum utility that go beyond claims of harmful interference. In such a situation, there would be an agreement between interfering systems to meet a certain statistical criterion for interference rise over noise, or a suitable availability requirement. This may be, for example, a stipulation that the Carrier-to-Interference ratio (C/I) will not exceed a threshold 75% of the time, and that one of the users will not face an outage for more than 0.1% of the time. This would be an improvement over using median characteristics or worst-case assumptions that cause the Interference to Noise ratio (I/N) levels to be set at fixed values. In addition, for these afore-mentioned cases, refined sharing mechanisms such as enhanced Licensed Spectrum Access (LSA) can enable the coexistence of the mobile service with other services.

- Introducing sensitivity analyses to evaluate possible different scenarios can lead to more realistic sharing assumptions. For example, IMT Active Antenna Systems (AAS) power weighting techniques, e.g., tapering, can be deployed in the mobile networks to reduce sidelobe levels and optimise performance. IMT AAS antenna models implementing such power weighting techniques could be used in sharing and compatibility studies, e.g., for sensitivity analyses. These studies should provide evidence that the effects of peak excursions of interferer levels will not cause unacceptable degradation to the performance of the incumbent services.
- Considering frequency ranges where sharing between the new service and incumbent services would require fewer restrictions to the new service, e.g., in terms of emitted power, could simplify the identification of candidate mobile/IMT bands. For example, restrictions could easily be simplified in the case of spectrum sharing with the uplink satellite service. For example, considering the frequency range 6-24 GHz and taking into account its worldwide frequency allocations, one candidate band for sharing is 12.75-13.25GHz (the uplink Ku-band). However, more bands suitable for IMT may exist, as it is expected that in some countries certain incumbent services are not in operation.

5.2 Initiatives to enable new spectrum for mobile/IMT: an overview

To further improve the B5G and 6G spectrum framework, adding new bands to the existing mid and high bands, as reflected in the resolutions comprising the agenda for the upcoming ITU-R World Radio Conferences (WRCs) in the time frame preceding 6G, will increase quality to most envisaged services. Considering the 2019 World Radio Conference (WRC-19), held in Sharm el-Sheikh, Egypt, from 28 October to 22 November 2019, relevant resolutions and recommendations, the need for several sharing and compatibility studies have been identified in preparation of the WRC-23 to ensure new frequency bands allocations to the mobile service on a primary basis. Resolution 811 from WRC-19 “Agenda for the 2023 world radiocommunication conference” [RR20] based on the results of WRC-19 and the requirements of existing and future services in frequency bands of interest, such as the 6425-7025 MHz frequency band in Region 1, as highlighted in the *considering* of and general characteristics can be found in Resoand in accordance with the related Resolution 245 (WRC-19). In addition, Resolution 812

"Preliminary agenda for the 2027 World Radiocommunication Conference" [RR20] highlights the importance of low bands as well, such as the sub-700 MHz frequency range.

In the resolutions mentioned above, several frequency bands are identified as prospective new bands for the mobile service. However, it is noted that still, some other frequency bands are of relevance for 6G, for example in the range 7-15 GHz and some bands in the the sub-Terahertz range (100-300 GHz).

It is worth noting that national regulatory initiatives are already in place in certain countries to enable new uses of sub-THz and THz spectrum. In the U.S., for example, the Federal Communications Commission (FCC) has made additional spectrum available for commercial uses across multiple spectrum bands between 95 GHz and 3 THz. A new category of ten-year experimental licenses for use of these frequencies has been introduced by the FCC since March 2019 when the Spectrum Horizons First Report and Order was adopted [FCC19].

According to [FCC19], device operation between 95 GHz and 3 THz is permitted with a maximum Effective Isotropic Radiated Power (EIRP) of 40 dBm (average) and 43 dBm (peak), measured with a detection bandwidth that encompasses the band of operation. Outdoor fixed point-to-point devices are also allowed to operate with a higher maximum EIRP of 82 dBm (average) and 85 dBm (peak), also measured with a detection bandwidth that encompasses the band of operation. In order to operate under the higher power limits, devices must utilise highly directional antennas with very narrow beamwidths (i.e., a minimum gain of 51 dBi, with a 2 dB reduction in the maximum permissible EIRP for each dB the antenna gain falls below 51 dBi) to ensure that the likelihood of harmful interference is minimised. Finally, equipment is not permitted to operate on satellites or onboard aircraft. The new experimental licensing framework allows licensees to commercially market equipment demonstrated using the experimental spectrum licenses.

In addition, the FCC has made a total of 21.2 gigahertz of the Spectrum Horizons bands available for unlicensed device use: the 116-123 GHz band, the 174.8-182 GHz band, 185-190 GHz band, and the 244-246 GHz band. Devices using these bands are allowed to operate on a non-interference basis while protecting both passive and active services. Reassessment and possible revisit of such spectrum allocations based on how uses develop is envisaged at a later date recognising that the frequencies above 95 GHz are potentially suitable for licensed use.

In the U.K., licenses for accessing 32.2 GHz of radio spectrum across four Extremely High Frequency (EHF) bands for wireless connectivity applications have been introduced by Ofcom since October 2020 [OFC21].

The EHF license grants access, on a 'one band per licence' basis and on a shared, uncoordinated basis, to any one of the following four bands: 57-71GHz, 116-122 GHz, 174.8-182 GHz, 185-190 GHz.

In the 57-71 GHz band, in addition to wireless access (e.g., small base stations fixed to a lamppost) or wireless backhaul (e.g., point to point links) solutions, to provide broadband services or help to connect other technologies such as IoT or M2M networks, lower power uses, such as high-performance wireless data, display, and audio applications, are envisaged as well.

At higher frequencies, in the 100-200 GHz range, the range of envisaged applications includes sensing, high resolution positioning, security systems, and high-speed data links.

Licenses are granted on a technology-neutral basis and can be used to deploy any device which meets the technical license conditions.

For systems operating in the 57-71 GHz band, the maximum permitted EIRP limit is 55 dBm and usage is restricted to fixed outdoor only. Systems operating at a maximum of 40 dBm EIRP may be covered by license exemption regulations.

Also, for systems operating in the frequency bands in the 100-200 GHz range, the maximum permitted EIRP is 55 dBm. However, outdoor use is permitted with additional power limits on EIRP depending on angles relative to the main beam in the elevation plane and on a per-band basis.

Licenses have an indefinite duration, but can be revoked (e.g., if the band usage regulation changes). A review of developments of future demand for services and applications using this spectrum is envisaged in 2024.

6 Sustainability

6.1 Sustainability KPIs and targets

The Hexa-X project has committed to achieving different targets related to multiple domains like performance, social value, or sustainability. Some of those targets are fully in the scope of "sustainability" and are mostly related to environmental impact, but some also concern areas addressed by other work packages of the Hexa-X project such as performance, automation, or frequency bands. For those, from a sustainability perspective, it is important to show how to consider their impact on the sustainability objectives of Hexa-X. The following section addresses the Hexa-X targets from a sustainability perspective.

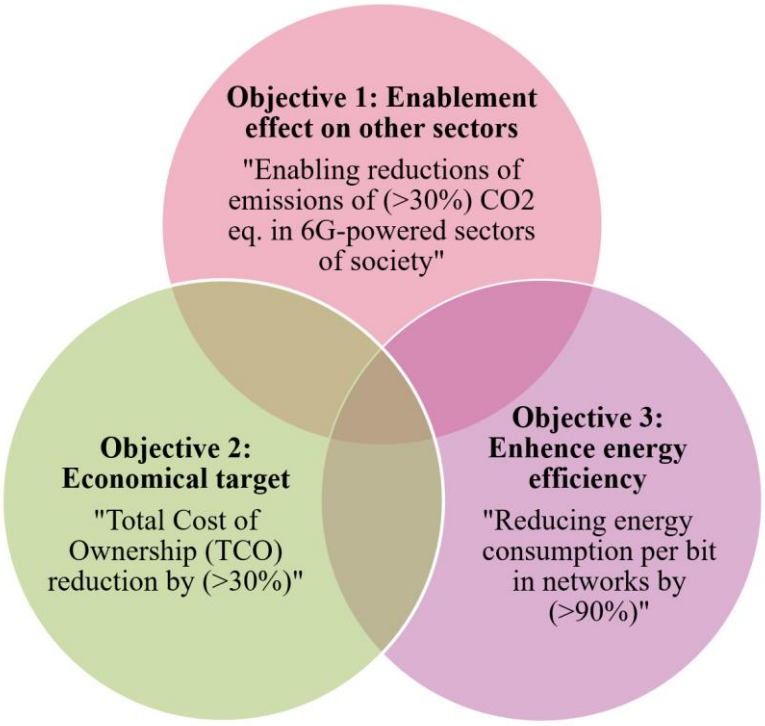

Figure 6-1: ***Hexa-X Work Package1 project objectives***

6.2 Addressing the targets

The following plan (see Table 6-1) has been elaborated to clarify the road towards each target. The reference scenarios and the scope are the main parts that still need to be defined.

Table 6-1: ***Plans toward Hexa-X Work Package 1 project objectives***

Objectives	Any need for clarification of the objective?	How to define the baseline?	Future plans to show this objective	Scenarios/ scopes considered for this objective	What are the difficulties to fulfilling this objective?
Enabling reductions of	Standardised methodologies	The baseline needs to	Providing initial high-	We have to define use cases for which	Establishing a reasonable

emissions of (>30%) CO2 eq. in 6G-powered sectors of society	(under development) are required to show the impact of ICT on other sectors. Moreover, the baseline has not been settled.	consider a scenario without 6G applied, hence it would create the basis for a comparison of 6G solutions vs non 6G solutions.	level methodology principles in D1.3. Possibly provide some quantitative scenarios by the end of the project (D1.4).	our proposed methodologies can be applied, in order to explore opportunities for decarbonisation of areas such as e.g., transportation, health, B2B, etc.	baseline demands an understanding of the future state of both 5G and 6G, as well as a way to define reasonable user scenarios Estimating the potential use and impact of future 6G solutions is challenging, especially since current methodologies builds on scaling measurements of actual application of technologies – an approach which is not applicable for future technologies. Consideration of the rebound effect needs a reasonable approach.
Total Cost of Ownership (TCO) reduction by (>30%)	It requires information or specialist on the TCO subject from all WPs in the project.	5G NR SA architecture with the new core network (5GC) and RAN (NG-RAN)	General proposal in D1.3. For D1.4 we require input from WP2-WP6 in order to be able to provide the output.	Select a few 6G features and enablers and compare them with a baseline to show the reduction in TCO.	We are not aware of specialists inside the project to analyse this.
Reducing energy consumption per bit in networks by (>90%)	**No, it is clear**	One possibility is to use 5G as a baseline	We will reflect on it in the D1.3. For D1.4 we can produce some quantities and provide them to technical WPs.	Methodologies are standardised in ETSI for calculating this KPI. However, since those methodologies addresses existing products and networks, their applicability for future technologies needs further consideration	Evaluating numbers on hardware and technologies that are not yet existing raises challenges regarding baseline and methodological approaches. Approaches from an earlier research

					projects (such as the European Earth project) should be evaluated to see if they could provide a way forward

In the following, we will give some details about how to achieve each target and the main challenges that need to be addressed.

- **Target 1: Societal impact target: enabling reductions of emissions of (>30% CO2 eq). in 6G-powered sectors of society**

Today, both the baseline and agreed detailed methods and harmonised standards that describe a clear methodology for evaluating the "enablement" impact of ICT on other sectors are lacking. Work is underway in ITU to provide assessment methods which are expected to become available during 2022. However, the traditional assessment approach is to refer to existing applications with a proven effect and scale those. Hence, assessment of future technologies will need to consider the fact that proven case studies will be lacking, implying an inevitable additional level of uncertainty and the need to adapt any existing methodology. It is also concluded that the overall effect of 6G (the aggregated effect of all potential, future use cases is beyond reach as the total use of 6G cannot be foreseen. Consequently, the evaluation of 6G can only be scenario based, and refer to specific use cases, mainly those defined by Hexa-X. Main challenges include the establishment of baselines, estimating impacts of future 6G solutions and their usage, and estimating the induced impact for a future scenario, potentially considering the rebound effect. To achieve this target, the following items still need clarification:

1. **Baseline**: probably a non 6G powered sector compared to a 6G-powered service. The main complexity is the data collection/estimation related to CO2 impact with and without 6G.
2. **Methods**: Section 6.4 on the enablement effect gives details on current international initiatives. Applying existing and developing methodologies to provide a transparent and well-founded result is a challenging task associated with significant uncertainties already for deployed technologies – to do that for future systems raises the bar even higher. Another difficulty will be to find a way to address rebound effects.
3. **Scope**: the enablement effect has to be defined and evaluated for specific use cases. Our ambition is to address one of the Hexa-X project use cases described in Section 2 (i.e., use cases) as a pilot.

The project is also discussing how to consider risk – use cases that could potentially increase carbon emissions and how to consider this from a methodological perspective.

- **Target 2: Total Cost of Ownership (TCO) reduction by (>30%)**

The TCO is composed of devices, networks, and datacenter operational expenses (OPEX) and capital expenses (Capex), both of these varying across use cases. The CAPEX estimation mainly depends on manufacturers product prices and many other factors that are challenging to estimate at this early stage of 6G development. The OPEX is country dependent especially for studies and benchmarks between developed countries and emerging ones.

Some published papers like [MLA11] give some perspectives based on MNO legacy networks. Considering sustainability and energy part, those studies estimated the "energy expenses" to account for about:

- 30% of Capex in the emerging market vs. 10 % in the developed market when including the energy supply and backup systems as well as the site technical environment (batteries, cooling system, AC/DC converters, etc.). Potentially, those parts could be neglected in the estimate of the 6G TCO solution, as most of these components are shared with legacy networks and not specific to 6G. However, future regulations may put stricter requirements which might suggest that these needs to be considered.
- The estimated energy OPEX is about 20% in the emerging markets vs. 10 %in developed markets of the total OPEX per site. This value represents essentially the energy OPEX used for powering the site. It is also important to consider also that 6G will represent only a fraction of this budget.

To achieve the target of –30% TCO, the following items have to be clarified:

- **Baseline**: focus on OPEX in the TCO dedicated to radio access networks. The OPEX split should be harmonized based on input from all Hexa-X partners. The Energy part is about 20% of the total RAN site OPEX [GSM19].
- **Methods**: Section 6.3 should help for energy assessment taking into account a products entire life cycle. We propose to focus on the expected energy consumption figures. However, solutions impacting the energy price have also to be considered, like using renewable sources or Power Purchase Agreement (PPA) to reduce the total energy cost. In general, smart grids are expected to be standard for 6G electricity supply.
- **Scope**: a dedicated use case should be defined.

It is noted that the outcome of this target is depending on the outcome of target 3.

- **Target 3: Reducing energy consumption per bit in networks by (>90%)**

This target is related to performance and efficient transmission as it links the energy consumption to the delivered data to users. To achieve this target, the following items need to be addressed and quantified:

- **Baseline**: The baseline should be existing networks including 5G at a defined point in time to be decided. Industry analysis and field measurements show that a factor of x10 has always been achieved at each technology change from 2G (40 kwh/Go), 3G (3 kWh/Go) to 4G (0.6 kWh/Go) as given by [ARCEP19]. 5G networks are designed to improve energy efficiency and their performance will develop over time. The following figure issued from an operator measurement on one specific product shows an example where 5G is four time more efficient than 4G (see figure) just considering spectral efficiency.
- **Methods**: Efficiency is basically expressed as the ratio between energy in kWh and e.g., data volumes in Gb evaluated during a time period. This is reflected in the MWh/Tb index given in ETSI TS 203-228 which has defined a methodology to evaluate this KPI at the network level. The applicability of this standard for future networks will be evaluated, as well as other possible intensities.
- **Scope**: The scope should focus entire network taking into account the usage phase.

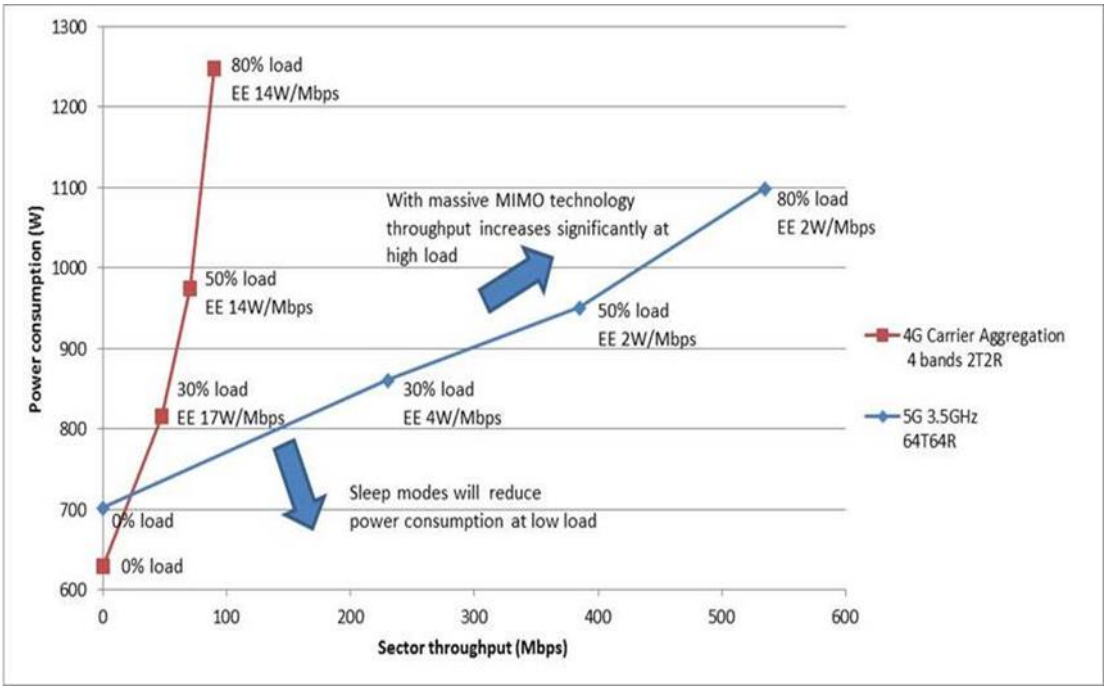

Figure 6-2: ***A comparison between 4G and 5G based on an operator measurement of one specific product***

6.3 Complementing priorities

To complement the 6G sustainability targets, it is fundamental to jointly into account all the sustainability aspects of networking, including hardware, planning, deployment, operations, and the entire equipment life cycle [HEX21-D51], starting from the standardisation phase. Indeed, the heterogeneity of resources and services, comprising communication, computing [HEX21-D41], control [HEX21-D71], sensing [HEX21-D31], etc., naturally calls for an ever-deeper E2E cross-layer design and optimisation and energy efficiency needs to be an integrated network design criterion [HEX21-D41], [HEX21-D61], [HEX21-D71].

While it is clear that mobile connectivity and networking could be a key enabler and accelerator of the decarbonisation transition, it is important that the GHG footprint and energy efficiency of mobile networks themselves be not ignored. The telecommunication sector is expected to serve 70 percent of the human population by the end of 2021 and 20 percent of the subscribers are expected to transition to 5G by 2025. Despite 5G being more energy efficient than LTE, the absolute energy consumption of 5G networks can be higher due to increased traffic density and data consumption unless built with precision and considering the co-optimization with older standards. This problem is further exacerbated by 5G deployments accelerating during the pandemic and the need to bridge the digital divide and expand access to the remaining 30 percent of the population.

In a survey of the mobile network operators and other key players in the telecommunication industry conducted by the GSMA, a significant majority of the respondents believed that energy efficiency is an important problem, and two thirds of the respondents believe that their energy costs will increase in the next three years. The telecommunication industry also has a roadmap to become carbon neutral by 2050 in line with the Paris climate accord, and over 30 percent of carriers have already made public commitments to have net-zero emissions by 2050.

AI-driven solutions are considered important for meeting these energy efficiency goals and emission targets. When asked about strategies to minimize their energy footprint, the stakeholders have the highest expectation in the transition to renewable energy, closely followed by AI. Other strategies include newer wireless technologies in 6G and decommissioning legacy networks. Of these strategies, AI is the most favourable since it is quick and cost effective to deploy, and is scalable with a favourable return on investment, unlike some of the alternatives.

Solutions that are expected to utilise AI include shutdown and sleep strategies based on user traffic patterns to reduce overall energy consumption. This is a very effective strategy since base-stations and RAN consume 70 percent of total energy in the network. Energy efficiency improvements can also be obtained through better resource allocation and load balancing, and predictive maintenance to reduce the number of site visits.

From the hardware point of view, the microelectronics R&D community is following to improve the energy efficiency of semiconductors technologies (neuromorphic chips, FD-SOI) and circuits architecture (active silicon interposers & chiplets, In-memory computing) with several orders of magnitude over the next decade to contain the digitalisation data deluge impact on energy consumption. In parallel, the microelectronics industry and all parts of the 6G supply chain should integrate its commitment towards sustainability, at all levels: Production (reduction of waste and water, reduction of critical materials), Eco-design of products (extended lifetime, life cycle analysis, choice of materials, enabling enhanced circularity), and End-of-life management (reuse and recycling strategies to reduce electronics waste). Except for new frequency bands HW, the partial re-use of previous generations of the infrastructure equipment could be considered, as well as a materials-optimising apportioning of functions, during the standardisation steps. At last, to favour the targeted 6G sobriety paradigm, the R&D will have to consider the challenge to maximize performance for a given resource at each sub-system level, without suboptimisation that risk the optimisation of an overall system and life cycle level. The link between hardware and software needs continuous attention, e.g., signalling structures are key to allowing sleeping of hardware which has a considerable impact on energy performance.

From an architectural point of view, cloud and service-based architectures, softwarisation, programmability, etc. are promising candidates to address 6G sustainability targets, thanks to their flexibility and adaptability to network changes [HEX21-D51], with the new architecture being able to accommodate new services (e.g., the ones involving massive computing), without increasing energy costs to beyond levels which are sustainable for businesses. To achieve that, energy related aspects should be taken into account at all phases, from planning [CDB21], [GAL21], to operations [SCA+19], [FRE21], [DDA+15], [TCD+14], also involving new protocols able to reduce the necessary control signalling, while still coping with dynamic network adaptability.

As technological enablers for energy efficiency in network operations, the exploitation of advanced sleep modes [SCA+19], [GFR20], MIMO muting [FRE21], low or zero-energy devices [HEX21-D71], [GSS+19], [MBM+20], etc., are promising ways to reduce the energy footprint, also achieving the "almost zero watts at zero load" target). To this end, AI/ML will help reducing complexity and, potentially, control signalling, thus avoiding energy waste [HEX21-D41], [HEX21-D71].

However, all these new technological enablers must be optimised online, with low complexity, and well-structured and possibly reduced control signalling, to avoid unnecessary additional energy consumption. To this end, a fundamental aspect of future wireless networks pertains to the exploitation of AI/ML, both as a flexible and low complexity tool for network optimisation and orchestration [BFY+21], and as an enabled application in computing as a service architecture at the networks' edge [HEX21-D41], [HEX21-D12]. AI-based orchestration can be used as a tool to increase energy efficiency, thanks to a proactive enabling and activation of resources, tailored to end users/verticals needs [HEX21-D61]. The use of AI is beneficial for energy efficient network optimisation, covering AI/ML-driven interface design, link level enablers (beamforming, adaptive channel estimation, coding, etc.), and system level enablers (mobility management, resource allocation, etc.) [HEX21-D41]

However, at the same time, future networks and services (such as edge AI in 6G) also call for joint optimisation of the supporting technologies (communication-computation-control co-design) [HEX21-D41], [HEX21-D71], by taking into account both communication (e.g., radio access, etc.) and computation (e.g., edge computing facilities) as sources of energy consumption. Therefore, energy efficiency optimisation should include the computing part of future networks, e.g., with the possibility of extending the concept of sleep operation modes to edge computing, thus adapting duty cycle of all network elements (AP, edge servers, etc.) to the current network conditions and traffic [WZY+18], [CHAN18], [MDD+21].

6.4 Defining the baseline for assessing the ICT environmental impact

This section will address methods and specifications that have to be considered for assessing sustainability targets. For any product or system, sustainability can be quantified and assessed with regards to different metrics for environmental impact like GHG emissions, water, and materials that are associated with the life cycle of those. In this section, we introduce methods, concepts, and definitions that are used for assessing the sustainability within our ecosystem. For this purpose, this section will give:

- an **overview** of the Scope concept which defines the organizational GHG emissions footprint, is here described to clarifying the role of different ICT actors in addressing emissions associated with designing and operating networks. Additionally, an introduction of the life cycle perspective is associated with understanding impacts at a product and system level.
- an introduction to the **life cycle assessment (**LCA) methodology used to estimate the life cycle impacts of systems for different environmental impacts.
- an overview of estimates of the current **ICT GHG emissions** which could potentially be considered when deriving the baseline for the 6G system; and
- some **trends** regarding the ICT sector development in relation to its GHG emissions.

6.4.1 Value chains, scopes, and life cycles – complementing perspectives

Climate change impacts can be analysed from both a company operation and value chain perspective, and with regards to the use and life cycle impacts of a product or service.

Coming from a company perspective, the first step towards decarbonisation, is to set a baseline which account for the GHG emissions from a company's direct emissions, known as Scope 1, from emissions related to its use of energy, known as Scope 2, and from all other emissions associated with the full value chain, known as Scope 3. These scopes definitions are defined by GHG Protocol [GHGP11] and referred to (by different terminology) by ISO 14064-2 [ISO18] and ITU L.1420 [ITU12].

More specifically, direct emissions under Scope 1 are related to direct emissions associated with the combustion of fossil fuels for company's cars, generator, heating systems for buildings, etc. Indirect emissions under Scope 2 are those related to the indirect emissions from the generation of purchased electricity, steam, heating, and cooling consumed by the company to carry out activities within its own organisation, office activities, inhouse manufacturing, IT, internal networks, tertiary buildings, etc. All remaining emissions associated with the value chain upstream and downstream are classified as Scope 3. Scope 3 accounts for emissions emerging from the manufacturing process of the goods a company sells or uses, as well as the emission related to those products end of life. It also includes electricity consumed by customers to power the products the company sells or leases, products transportation, franchises, etc. Moreover, it covers employees' commuting and business trips. The boundary between Scope 1 and 2 on one hand, and Scope 3 on the other hand, depends on the organisation of the company. For instance, depending on who owns the car that delivers a product those emissions may be categorized as either Scope 1 or Scope 3.

The definition of scopes is complex as Scope 3 of one company is composed by the Scopes 1 and 2 of other companies in the ecosystem. Hence, the same emissions are belonging to different scopes depending on where you are in a value chain. Ultimately, only Scope 1 emissions can be aggregated without any risk for double counting while other emissions represent emissions associated with the company but occurring in other parts of its value chain.

The other lens to apply when analysing climate change impacts is that of the life cycle assessment of the products or services delivered by an organization. The method applied in this case is called LCA and is further detailed in Section 1.2.

Applying the product life cycle perspective or the company carbon footprint reporting perspective is different, but the two approaches are closely related: A company carbon footprint should consider both, the use of products (Scope 3 downstream) and manufacturing, etc (Scope 3 upstream). At the same time, an LCA of a product needs to consider also the supporting activities of the involved companies.

Figure 6-3 seeks to illustrate these connections and shows the allocation of different emissions both from an LCA and organisational perspectives. Please notice that, following the LCA standard [ITU14], for products and systems, emissions are grouped as material acquisition, production, use and end-of-life treatment processes. Then, for a given product, emissions of these processes will contribute to the scopes 1, 2 or, 3 of the providers (blue squares) and of the operators (pink squares).

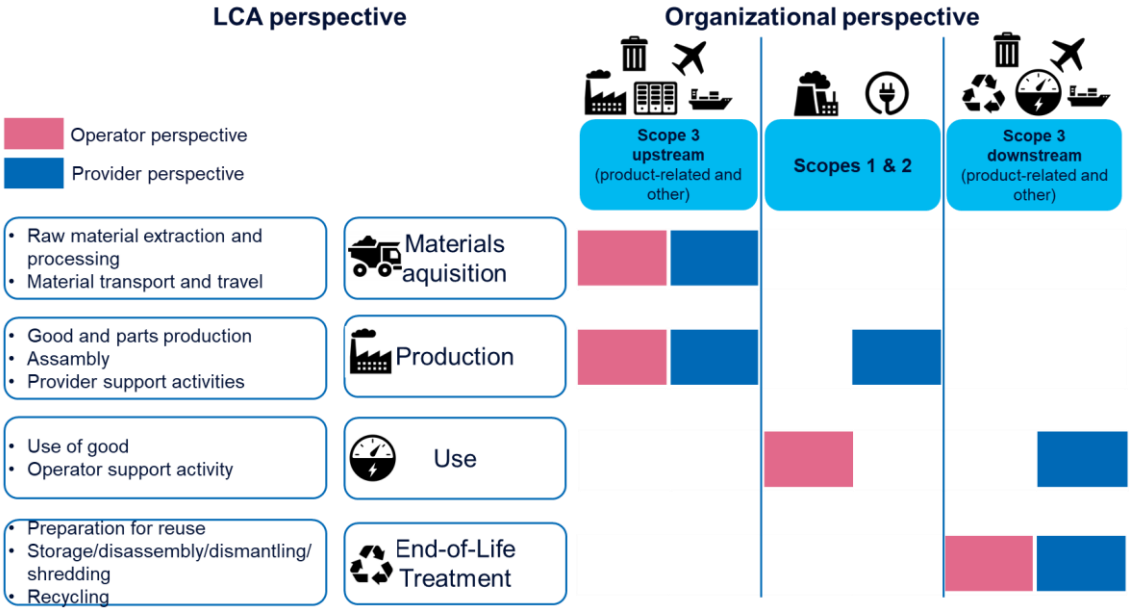

Figure 6-3: ***GHG emissions of a product from the perspective of an LCA and organisational carbon footprint view***

In addition to scope 1,2, and 3 introduced above, the term ‘Scope 4’ is sometimes used (other terms with similar meaning include ‘handprint’, ‘enablement’ (used in other sections of this document e.g., 6.5), and ‘avoided emissions’). Scope 4 refers to those emissions occurring outside of a company´s value chain but due to the impacts associated with the use of its products. A typical example is the use of video meetings replacing physical ones thereby providing the opportunity to reduce emissions associated with travelling. Scope 4 may represent much larger effects than Scope 1, 2 and, 3 emissions and may be either positive or negative. However, scope 4 emissions are considered additional from a reporting perspective and are not considered to be part of a company´s carbon footprint. Currently, Scope 4 is less well defined than the other scopes [AHT+20], [BVM+20], [CBM+20] and more accurate assessment methodologies are under development, e.g., in the frame of ITU-T SG5.

This section refers mainly to the different perspectives that apply when analysing climate change impacts as the scopes are defined in relation to a company´s carbon emissions. However, LCA and a company´s organisational environmental footprint should also consider a broader range of effects such as resource depletion, toxicity, oceans acidification, etc. Typically, an analysis of only carbon emissions cannot be used to predict other environmental impacts.

From the perspective of 6G research, the important task is to develop the 6G system in a way that minimizes the environmental life cycle impacts of the system (Sustainable 6G) while at the same time developing use cases and capabilities that maximize its positive effects and suppress any negative effects (6G for Sustainability). Realising this, future use of 6G could also help companies to reduce their emissions across the different scopes.

6.4.2 Assessing environmental footprints

To evaluate the environmental footprint of an equipment, a network, or a service, LCA provides a comprehensive method for capturing its overall impact. Indeed, LCA covers the entire life cycle of the studied system (from raw materials extraction to end-of-life treatment) and is able to deliver results for multiple environmental impact indicators (e.g., climate change, abiotic resources depletion, and water consumption). The LCA method has been standardised for all type of system by the ISO 14040/14044 series [ISO06a] [ISO06b], and specific recommendations for the ICT sector has been published by the ITU-T [ITU14] and the ETSI [ETS14]. At the European level, the Product Environmental Footprint [PEF] program has also been working on harmonisation of LCA methods in order to deliver more consistent and transparent results and communication to consumers.

Carrying out an LCA is, in a few words, a matter of first performing a Life Cycle Inventory (LCI) by collecting product and process data (kg of raw material, km of transport with a lorry, kWh of electricity, etc.) to cover all the input and output flows at the studied system level. Figure 6-4 shows the different activities that have to be considered according to [ITU14].

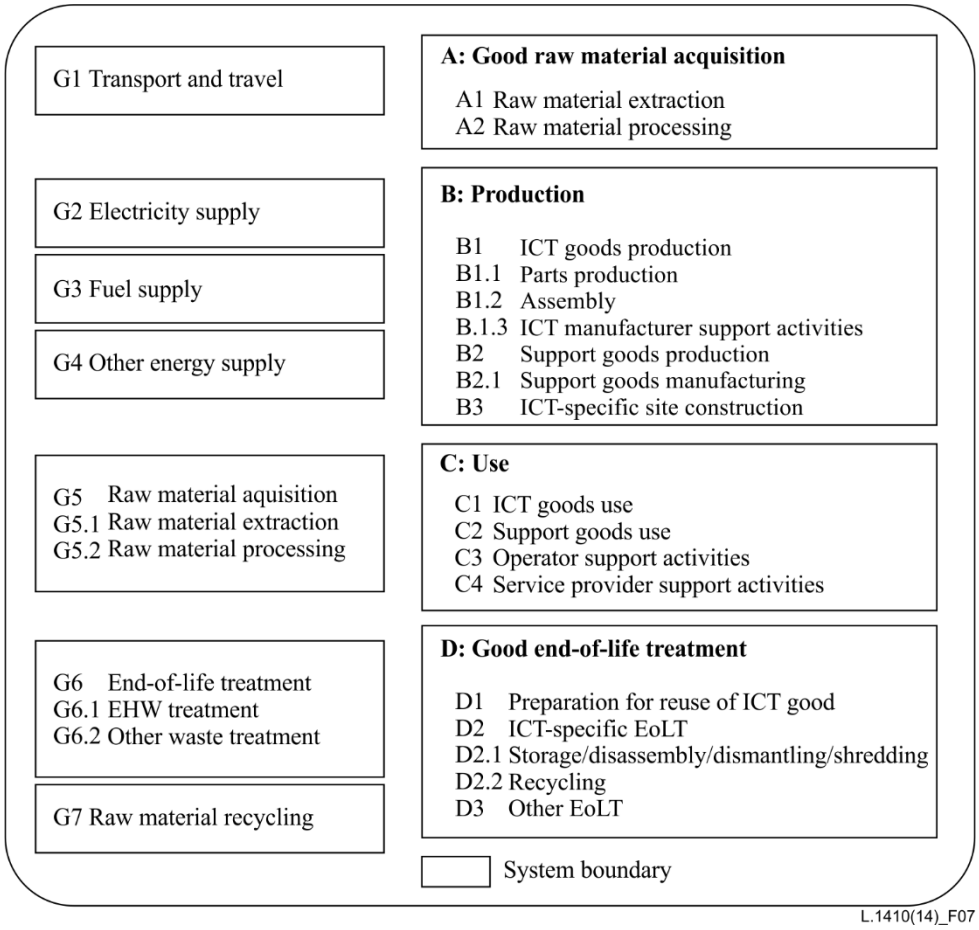

Figure 6-4: ***The system boundary of the product system for LCAs of ICT goods, networks or services (from [ITU14])***

Ideally, the collected data should emerge from primary sources (i.e., actual measurements carried out for the studied system, like load rates and energy consumption measured on an extended period of time) - not from a secondary source (i.e., generic data such as the theoretically calculated average energy consumption of a base station at specific load levels). However, in reality most LCA typically contain a combination of primary and secondary data which is acknowledged by the ITU standard [ITU14]. A dedicated LCA software (e.g., Gabi, SimaPro, EIME) and its life cycle inventory database may provide some of these secondary data and is also used for converting all the collected data into environmental

impacts during the Life Cycle Impact Assessment (LCIA) phase. Based on this, the interpretation phase evaluates and contextualizes the results, which is a key activity of any LCA study.

LCA relies on actual data which makes it hard to perform LCA studies of 6G products and systems before their standardisation, development, deployment. Still, for a project such as Hexa-X, LCA could offer many other possibilities. For instance, as 6G will require critical raw materials such as Gallium for GaN based semiconductors, LCA can be used to provide an early assessment on the potential use at a large scale of such semiconductors within the framework of an entirely new mobile radio network technology. Research questions to investigate may include, for instance: Will the additional use of Gallium consumption driven by the manufacturing of 6G equipment have a significant impact on abiotic resources depletion (and thus on available reserves)?

LCA estimates can also be developed by modelling different changes compared to today´s features and systems and estimating the associated change in environmental footprint for a future 6G scenario and interpreting this in relation to a set of net-zero carbon strategies. In other words, assuming that by 2040 telecommunication operators have to be carbon neutral, how could the overall system be defined and how should the different hardware and software pieces of 6G contribute to this target, under different systemisation and design choices? Such studies could be helpful, but measures should be taken to avoid suboptimisation whenever focusing on one or a few pieces of an entire system.

6.4.3 The current estimates of the GHG emissions of the ICT sector

Comparing LCA studies is complex and ITU L.1410/ETSI EE ES 203 199 [ITU14], [ETS14] makes clear that results from different studies cannot be compared directly but demands a detailed understanding of the associated models, assumptions, and data sets already at an equipment, network or service level. Addressing the additional complexities of a sector level assessment, ITU L.1450 establishes a standard for estimating carbon emissions of the ICT sector, and assessments based on this standard could serve as a basis for understanding how the GHG emissions evolve at a sector level.

In the recent years, several studies on the environmental impact of the ICT sector were published ([MAL18], [AND20], [BEL+18], [MAL18a]). The considered assumptions and modelling vary from study to study. To be able to make an apples-to-apples comparison, we first look at the different components of ICT. Based on ITU-T L.1450 [ITU18] the ICT sector (seen from the perspective of its deliverables) is composed of networks, user devices, and datacentres. Operator networks, specifically, are composed of exchanges, BSs, routers, etc., user equipment includes devices such as smartphones, PCs, IoT devices, and so on. Data centres include Over The Top (OTT) servers. In comparison, ‘the digital sector’ is a broader concept which has not been standardised from an LCA perspective. However, it typically includes, alongside ICT, user devices traditionally associated with the Entertainment & Media (E&M) sector such as TVs and TV peripherals which, according to the standard, fall outside the boundaries of the ICT sector.

The above-mentioned studies differ in terms of scope as some consider only the ICT sector (as defined by the standard) and others consider the less clearly defined concept of the “digital sector”. To make meaningful comparisons, we disregarded E&M devices as not all studies included them and focus on ICT. The carbon footprint of the ICT sector (without TVs) is estimated in 2020 to make up for 1.1-2.1% of global GHG emissions [FRE+21], see Figure 6-4 below. The different levels of recentness, availability, and transparency of data, as well as the degree of primary data and data age used by the authors of the studies, explain the variations in the estimates. In particular, [MAL18a], completed by [MAL18], which is one of the most comprehensive and recent studies, provides an estimate of 1,4% of overall global GHG emissions for 2015 (1.9% if TVs are included), based on a large sample of measured data as reported by companies and estimates that that level will stay quite stable until 2020 (an estimate which is confirmed by ongoing research following the same methodology) – this level has also been agreed as a sector baseline for the decarbonisation trajectories developed jointly by ITU, GSMA and GESI and applied by SBTi [ITU20]. Of this 1.4%, user devices represent 54% of emissions, networks 25%, and data centres 22% (resulting in 101% overall due to round off effects). From the global perspective, the embodied/use stage emissions ratio is closer to 50/50 for aggregated user devices due to short life spans and less intensive usage, compared to network and data centre equipment for which

use stage emissions represent the absolute majority of emissions due to longer life spans and as the equipment is used 24/7. This balance looks different in countries with low carbon electricity mixes, and an operator´s purchase of renewable electricity increases. In both cases the relative importance of embodied emissions increases. Moreover, assessing a wider range of impact categories typically shows that other life cycle stages than usage are dominating for categories such as toxicity.

6.4.4 Trends in GHG emissions of the ICT sector

There are three major drivers for the future trends in emissions of the ICT sector: the decoupling between data traffic and energy consumption, the growth of ICT energy consumption, and the decarbonisation of ICT energy.

Data traffic has been significantly increasing for decades, and energy consumption has fortunately not followed the same curve. This is due to energy efficiency gains taking place everywhere in the ICT systems. In operator networks, the energy efficiency (in GB/kWh) is multiplied by 10 from 4G to 5G. Between 2010 and 2018, the number of data centre compute instances has been multiplied by a factor of 6.5 [MSL+20] while the operation energy consumption of data centres increased only marginally by 6% in the same time period.

Despite energy efficiency gains, the growth of the sector is continuing, and the energy consumption of the sector has increased nonetheless and is expected to continue to do so, though moderately. The rise of new uses of ICT (blockchain, IoT, AI…) calls for vigilance regarding energy consumption development of the sector. On the other hand, a saturation of several market segments is likely to incur a slowdown in the growth of energy consumption. Modernization of both fixed and mobile networks also plays an important role. A key observation is that the development may take different paths depending on customer preferences, suppliers´ sustainability ambitions, legal frameworks, etc. A normative but plausible trajectory developed jointly by ITU, GSMA, GESI, and SBTi put forward a normative scenario where electricity is allowed to grow, but only to the extent that it keeps within its own relative share of emissions while emission levels shall reduce by 45% 2020 to 2030 [ITU20].

Higher energy consumption does not necessarily translate to more emissions, as actors of the ICT sector move to less carbon-intensive sources of energy. This can be seen in Figure 6-5, which depicts projected 2020-2030 trends for data traffic (blue), electricity consumption (yellow), and carbon footprint (grey). As demonstrated in the L.1470 trajectories a majority of the sector´s emissions are associated with the use of electricity, also when considering manufacturing, etc, implying that the use of renewables is the most important measure to keep down carbon emissions. This materialises through an increase in renewable energy supply. Indeed, while the global electricity mix has an emission factor of around 0.6 kgCO2e/kWh, renewable energy has a factor of ca. 0.1 kgCO2e/kWh. Therefore, switching entirely to renewables throughout the life cycle would reduce energy emissions of ICT by 86% [FBW+20], while the ITU scenario implies this number to be slightly lower, around 80%. Note that the ICT sector uses a large share of renewable energy which, on the one hand, could be said to limit the ability of other sectors to use renewables, as the worldwide supply of renewables is still low. However, the competition for renewables would also mean that it would be more attractive to invest in renewables, hence providing a positive market push. In any case, as long as the access to clean energy is insufficient, all sectors need to be mindful about their energy performance. This is also a cost issue. Thus, switching to renewables would not be sufficient but must go together with actions to control overall energy consumption. Additionally, to address emissions associated with use of chemicals and materials complementary measures are needed.

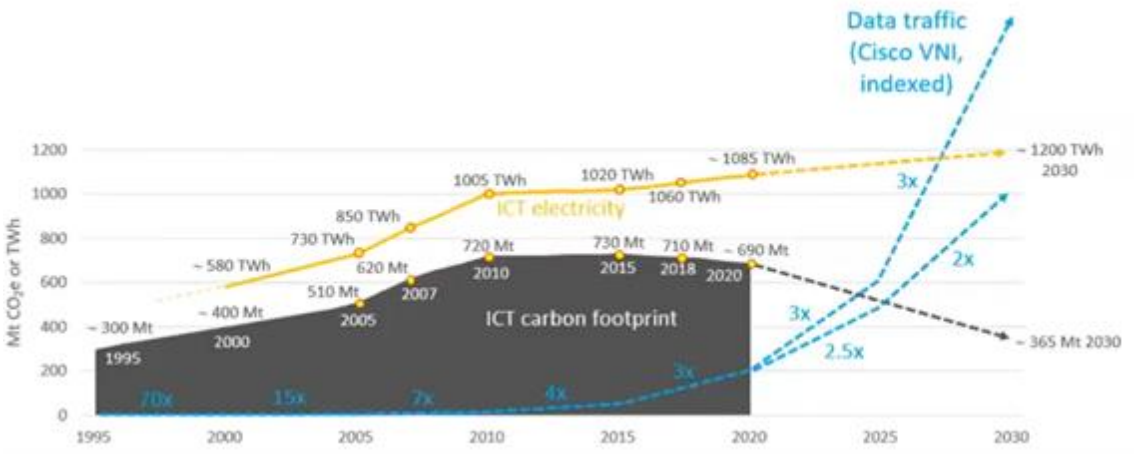

Figure 6-5: ***Trends in carbon footprint, data traffic and electricity consumption (Source: FBW+20)***

The magnitude of each driver in the future is uncertain at this point, so there is no consensus on the future trends in emissions of the ICT sector. It is also important to remember that the future trend is the consequence of our actions, thus impactable. The authors of [MAL18] indicate, based on ongoing research, that the emissions keep stable 2015-2020 and put forward a feasible and necessary reduction by around 50% by 2030 in line with the normative L.1470 scenario [ITU20]. Let us recall that the L.1470 demand that the sector, to help stay beneath the 1.5°C warming limit, should reduce its emissions by 45% in 2030, towards net-zero by 2050 to follow a 1.5C ambition [ITU20]. Some years back it was a common position that ICT should be allowed to reduce its emissions at a slower pace, given that it makes it possible for other sectors of the economy to reduce their own emissions ('enablement effect'). However, at this point, such a view is less common, and the majority of actors seem to have accepted that avoided emissions cannot be used as an excuse for not addressing its own footprint – especially since the ICT sector is considered as an "easy to abate sector". Nowadays, footprint reductions and avoided emissions are normally seen as complementary.

6.5 Enablement effect: How 6G could help

Hexa-X (and other initiatives working on the 6G definition) have placed environmental sustainability as one of the key driving forces towards 6G. 6G should be sustainable ("Sustainable 6G") and, moreover, 6G should be capable of supporting services that will contribute to sustainability: "6G for sustainability" (commonly referred to as the "Enablement effect", in particular in relation to climate impacts).

In order to deliver towards this goal, 6G will need to be efficient, from the performance and cost perspectives, in the delivery of diverse applications and focus on those applications that bring possible positive sustainability effects. Wireless in general, and 6G in particular, can play a role in areas like: environment monitoring and protection, protection of humans, utility management (e.g., water network protection and optimisation, and maximization of use of renewable energy), optimisation of production (e.g., circular production systems, optimisation of logistics and manufacturing facilities, and precision agriculture), health and security, etc. However, the evaluation of the enablement impact should follow rules and guidelines that are commonly and widely adopted and standardised. This section will give some highlights about ongoing methodology development in this new domain that will help assess the target of a 30% enablement given this project.

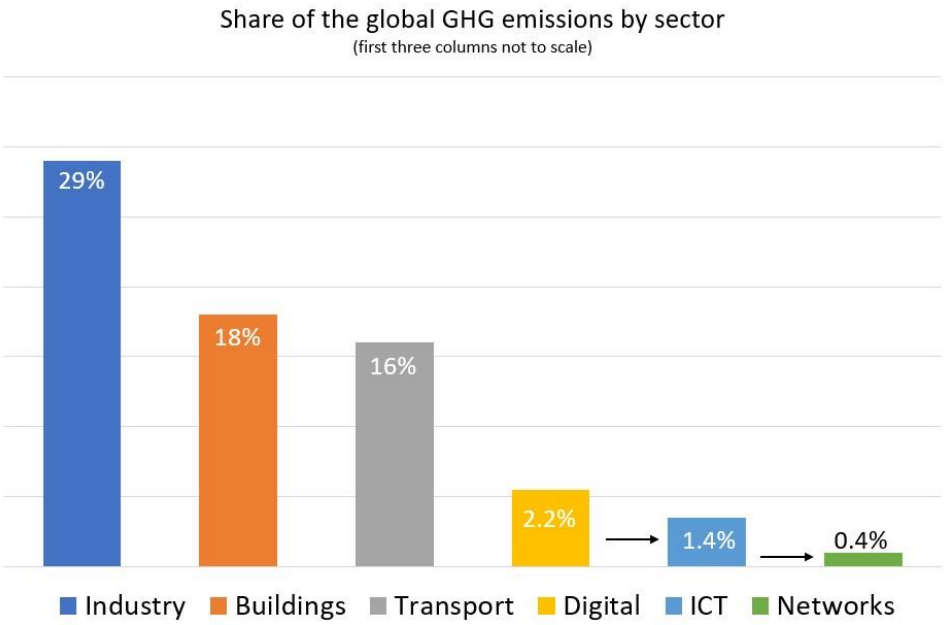

Figure 6-6: ***Share of the global GHG emissions for several major sectors (source RIT21), for the digital/ICT sectors and for the operator networks within ICT.***

As mentioned in Section 6.4.1, the enablement effect has several names including also avoided emissions and handprint, or – when referring to both negative and positive effects – induced effects. The following section will stick to the term 'enablement' for consistency although different initiatives may use different terms.

6.5.1 Assessing the enablement effect

The enablement effect is most commonly associated with solutions that could help avoid GHG emissions. This effect has been put forward since long and early estimates were put forward by e.g., GESI and WWF. However, this section is focusing on the more recent work by The French Agency for Environment and Energy Management (ADEME) [ADE], GSMA supported by Carbon Trust [GSM] and Bieser et al. [BSH+20] A common basis for most work in this area is the definition of a baseline scenario, a definition of a scenario with a solution that reduces GHG emissions applied and a comparison between the two. This is by necessity a hypothetical scenario since the two scenarios cannot exist at the same time.

The ADEME gives methodological guidelines for the quantification of actions which helps avoid emissions [ADE]. Avoided emissions correspond to reductions in emissions realised by ICT products and/or services but take place in other sectors than ICT. To quantify the emissions avoided, say, by a low-carbon solution permitted by ICT, there must be a reference scenario against which the low-carbon solution is compared. This scenario must be described and justified, and its hypotheses explained. Given the complexity of quantifying avoided emissions, it is important that a critical review be conducted (as put forward by ISO 14044 for instance), to make the estimation of avoided emissions credible. To give an example, telemetering is a low-carbon solution permitted by ICT, and a reference scenario to quantify the avoided emissions of telemetering could consist of a technician going on premises X times a month and driving Y kms with a car emitting Z kg CO2/km.

Avoided emissions permitted by ICT have been studied for quite a few years and several assessments have been put forward. Generally speaking, the studies undertaken seem quite often to be fragmented as they rely on selected use cases and sometimes refer to specific regions. Moreover, considering that ICT is pervading almost any sector, the use case selection may appear arbitrary. In particular, studies are often accused to be cherry picking by only considering usage with an expected positive effect.

Before the Covid outbreak, the GSMA together with Carbon Trust developed the white paper "The Enablement Effect" [GSM]. In that study, the enablement impact of mobile communication was

estimated to exceed 2 billion tons of CO_2 avoided in 2018, almost ten times greater than the total CO2 emissions of mobile networks globally. In a previous study in 2015, this enabling ratio was estimated at around 5:1. It should be noted that such ratios are advised against by the WRI that postulates that such comparisons should include all services and use cases of a company or sector – also negative ones.

The GSMA & Carbon Trust study adopted a rather straightforward calculation process for the avoided emissions (in kgCO2e) by multiplying avoided emissions factor by a quantity metric (e.g., number of IoT connections or number of smartphone users). One interesting characteristic of the study is that it combines measures of impacts with a wide behavioural study to understand the usage scenario. However, not all the findings are detailed in the documents, and there are some assumptions that could need further discussions, and such results could also be outdated due to the pandemic. The study also outlines some methodological considerations that are applicable also to other approaches and are therefore useful to understand the main challenges that should be addressed:

- Business-as-usual base case: it can be a challenge to determine what the base case would have been in the absence of the new solution.
- Allocation of avoided emissions between relevant technologies: any enablement solution typically relies on different technologies (i.e., not just telecommunication) and the identification of a consistent allocation is challenging, However, in line with scope 3 it may not be necessary to allocate emissions to a level that make emissions uniquely associated with one estimate only – but at the very least it would be important to make clear what technologies are contributing and in what way.
- Rebound effects need to be considered. This is especially the case for a direct rebound as that may impact already modelling of second order effects (see next section).
- Analysis uncertainties and completeness: on the one hand, data and assumptions might be affected by uncertainties and consistency issues; on the other hand, it is not possible to perform a comprehensive analysis and it may be challenging to correctly identify the most relevant scenarios/use cases to be taken into account

These considerations are in line with [BSH+20] and [BVM+20], [CBM+20].

Some methodological flaws regarding the GSMA & Carbon Trust study from 2019 have been suggested [ROU21], such as low data reliability, ultra-optimistic hypotheses and estimations, and the fact that it is unclear whether the advertised “avoided emissions” occurred in the first place. One example of an ultra-optimistic hypothesis is related to the use case “Accommodation sharing”. This use case alone accounts for 221.5 MtCO2e avoided – more than the global emissions of mobile networks – but it relies on an assumption sourced from a 2014 study from a sharing platform company, which claims that the emissions of a residential room are 89% lower than a hotel room. This claim is not justified and seems excessive. For the sake of comparison, the difference in energy consumption between hotels and residential units is approximately 20% on average in France and not 89%. Using this new factor would result in avoided emissions divided by 4 for this use case. This specific criticism is in line with the more principal observation on pitfalls put forward by [BVM+20], [CBM+20].

More recently, the GSMA and the carbon trust have released a new whitepaper with a high-level quantification of emission avoidance, titled "Industry pathways to net zero: Mobile and digital technology in support of industry decarbonization" [GSM21].

In this whitepaper, the total annual GreenHouse Gas (GHG) emissions from four sectors, namely manufacturing, power, and energy generation and distribution, transport, and buildings are calculated to be 52.5 gigatons. These four sectors collectively contribute to 80 percent of the overall GHG emissions. An overall reduction in annual GHG emissions of about 27.4 gigatons (approximately 50% of current annual emissions) across these four sectors is expected by the end of this decade to remain in line with the net-zero emissions target by 2050.

Mobile connectivity and networking are expected to improve the overall energy efficiency in all four sectors. In the base case, advances in mobile connectivity and networking are expected to enable avoiding 10.3 gigatons of GHG emissions by 2030, which collectively is about 27 percent of the overall

savings target. Decarbonisation in the manufacturing sector is primarily driven by productivity and energy saving yields through higher grade automated production equipment with IoT sensors that can be monitored and adjusted in real time by machine learning algorithms in the cloud. In the power and energy sector, the analysis covers emissions the sector can directly influence, which are those from the industry's own activities in extraction and refinement of fossil fuels, as well as distribution and electricity produced for use by other sectors. The reduction in carbon emissions comes from efficient distribution through connected grids, and smarter consumption of energy by customers through smart metering. In the transport sector, the IoT systems and onboard connected telematics are expected to improve the charging and fuel efficiency, respectively, of electric vehicles. GHG emission savings are also expected from a larger portion of the workforce working from home. Urban traffic management systems that use IoT and wireless connectivity can help reduce congestion and emissions from idling. Lastly, IoT sensors are expected to bring about the majority of the energy savings in buildings through smart electricity and gas meters, and occupancy monitors. Since the approach is similar to that used in "The Enablement Effect" a more in-depth review would be needed to understand the applicability of this approach.

Other approaches have been proposed in literature like the study forecasting the effects in 2030 Switzerland [BSH+20] concluding that the CO2 reduction potential outweighs the total footprint of the mobile infrastructure (even if only a few selected 5G applications were chosen).

In addition to reducing GHG emissions and mitigating climate change, digitisation and connectivity also have socioeconomic benefits, namely, improved public health from cleaner air, economic diversification by creating jobs and expanding access to them, increased access to public services, and civic engagement, and higher levels of productivity. These benefits shall not be forgotten although they are harder to quantify as they are also important pillars of a sustainable society.

6.5.2 Methodology for assessing the "Enablement effect"

ITU [ITU14] refers to three levels of environmental effects associated with its products and services (see Figure 6-7):

- First order effects: Often referred to as the footprint
- Second order effects: Also known as enabling effects, avoided emissions, etc (if positive), a more neutral term is induced effects
- Other effects; Rebound and other behavioural and societal effects,

While LCAs of ICT´s first order effects are well established albeit complex, and covered by standards, ICT for sustainability is a less mature area. ITU-T L.1410 part II [ITU14] provides a first attempt in the direction of comparative LCA and sets some fundamental principles. However, it is not considering several of the complexities around such assessments [CBM+20], [BVM+20]. This includes such as the need to improve the goal and scope definition (analysing differences between ICT substitution and optimization, the time perspective of the assessment, the challenge of a hypothetical baseline for the situation without the ICT solution, and the differences between modelling and case studies) as well as the often-ignored influence of rebound effects and the difficult extrapolation from case studies to larger populations. Moreover, L.1410 [ITU14] is not giving guidance on the assessment of multiple services, or assessments performed from an organisational perspective. For this reason, ITU-T has now started a work item that will provide more elaborated guidance on the assessment of the induced effects of ICT which is planned to be finalized during 2022.

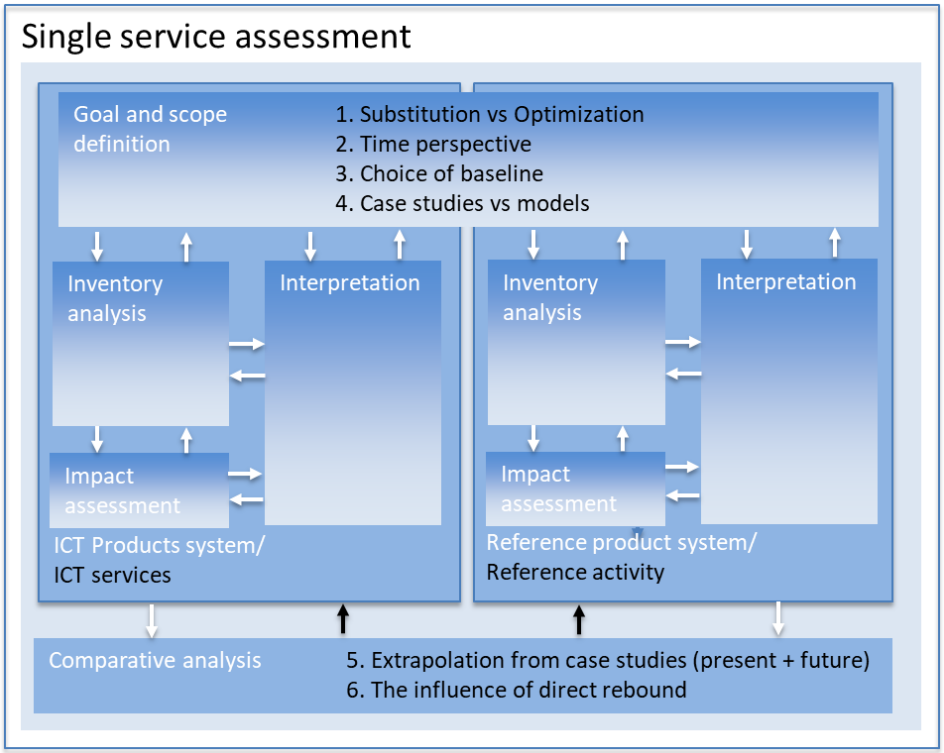

Figure 6-7: ***Improvement areas for assessment of enablement in [ITU12] and [ETS14] (Source: [COR20])***

The ongoing work of ITU will look into other methodologies as applied by different studies and seek to build on the learnings of each including some of those mentioned in Section 6.2 and others including aspects of ADEME QuantiGES, GSMA Enablement Effect, WRI avoided emissions, etc. However, a common characteristic for all those is that they build on access to actual data. This means those will not be directly applicable to estimate the value of future technologies. Hence none of these methodological approaches can be applied for 6G without modification.

6.5.3 Need to link with Hexa-X UN-SDGs

This chapter has mainly focused on environmental sustainability, in a particular climate. However, 6G can offer vast potentials enablement in associated with a broader range of environmental and other sustainability enablement aspects; whether it is in connection to greenhouse gas emissions, to addressing poverty, hunger or health, education, or any other societal area as represented by the UN sustainable development goals (UN SDGs). The UN SDG framework consists of 17 goals with 169 targets whose achievement is regularly monitored through 232 unique indicators. The UN SDG framework names 7 ICT specific indicators that characterise ICT skills and the usage of the Internet, computers, mobile phones and networks, and fixed broadband but it does not provide a clear mapping to mobile communications or design criteria for 6G. Studies have developed a link between mobile communications and all 17 UN SDGs [GSM18] and a connection between 6G and UN SDGs was provided in [MAA+20].

The D1.2 deliverable outlined how the Hexa-x use cases can be mapped to a delivered value for which an association can be made to one or several UN SDGs on target level, or alternatively to other values, e.g., ease of life or entertainment. In D1.2 deployment descriptions, one of the use cases was further described to map to a number of SDG targets, i.e., the e-health for all use cases. In this deliverable, a similar analysis is done for additional Hexa-x use-cases in Section 6.3.

Key value indicators for 6G are related to the assumed (hence at this stage hypothetical or theoretical) enabling effect of the use of the service. The beneficial impact for humans and society – the actual value - cannot be assessed before the commissioning of the 6G network with the specific applications needed for the different use case utilisations. The actual impact will depend on much more than just the

availability and performance of service including policy and financial aspects. However, within the sphere of control of the ICT industry, the task is to provide the infrastructure needed to deliver a trustworthy and good quality of relevant services at a relevant cost, covering sufficient areas and accessible for those needing it. The value quantification is thus heavily deployment dependent and at this early research stage, the focus lies rather on identifying important capabilities the network must have to allow realisation of a use case or, applications and needed deployments that can deliver at scale on the SDG targets. This understanding will support in forming requirements on 6G to realise sustainability enablement.

7 Next Steps

As a next step the content of this report will be updated and extended in the following way:

Task 1.1 on Common Vision will become active again for the last 6 months of the project in order to update the Hexa-X Vision on 6G based on the work inside and outside of Hexa-X.

Towards D1.4, the focus of the performance and KVI task will be to deepen the analysis of the technical enablers proposed in other Hexa-X work packages regarding their potential to contribute to the key values sustainability, trustworthiness, inclusiveness, and flexibility. This includes the impact of the architectural concepts proposed in this deliverable on performance indicators and limitations in potential deployments. Regarding the key values trustworthiness and sustainability, collaboration with the Security/Trustworthiness and the Sustainability tasks will be continued and deepened.

This report contained the gap analysis performed by the architecture team on the various technical enablers. The intention was to identify the requirements from the architecture as well as the possible impact each technical enabler will have on the next generation 6G architecture. Moving toward D1.4, the team aims to conduct a comprehensive study and eventually provide a concrete conclusion on the design of the 6G E2E architecture. In particular, by deepening the knowledge on each technical enabler and collaborating closely with technical work packages to identify the details of each building block of future mobile network generation.

The security team is planning to focus the future T1.7 activity on a deeper understanding of the relevant security enablers and their mapping on the reference E2E architecture, adapting this mapping as the architecture is refined, and on continuing collaboration with the technical WPs in aspects such as privacy implications, the applicability of physical layer security techniques, threat models, management frameworks, and the implications of the pervasive use of AI. In addition, there are several topics that have recently emerged that would require more detailed consideration, in particular the mechanisms for LoT assessment and its implicit subjective component, and the conditions for providing Security-as-a-Service to external networks. Finally, the team is considering the possibility of a dedicated security analysis of specific use cases, in order to identify the 6G-specific *security delta* for them.

Concerning the Sustainability task, over the coming months, the focus will be mainly on the methods and roadmap for achieving the three associated KPIs namely enabling reduction of emission of CO2 equivalent in 6G-powerd sectors of the society, reduction of the total cost of ownership, and reduction of energy consumption per bit in the network. For each KPIs the team defined the baseline scenario, the boundaries for the evaluation, and the main method for achieving them. For this reason, the action plan of the sustainability task will be shared with other work packages, especially those contributing to the same KPIs. Additionally, harmonizing the methodologies conducted in this project with international and standardised methods are in the plan. This step is necessary as adaptations may be needed.

The results of the described work will be reported in Deliverable D1.4 and published on the Hexa-X webpage.

8 References

[22.071] 3GPP TS 22.071 “Location Services (LCS); Service description; Stage 1”, Release 16, v16.0.0., July 2020.

[22.104] 3GPP TS 22.104 “Service requirements for cyber-physical control applications in vertical domains,” Release 16, v16.5.0, July 2020.

[22.261] 3GPP TS 22.261 “Service requirements for the 5G system,” Release 16, v16.16.0, December 2021.

[22.804] 3GPP TS 22.804 “Study on Communication for Automation in Vertical domains,” Release 16, v16.3.0, July 2020.

[23.288] 3GPP TS 23.288, “Architecture enhancements for 5G System (5GS) to support network data analytics services," Release 17, v17.3.0, Dec 2021.

[23.501] 3GPP TS 23.501, “System architecture for the 5G System (5GS)” Release 17, v17.3.0, Dec 2021.

[23.502] TS 23.502, “Procedures for the 5G System (5GS); Stage 2,” Release 17, v17.3.0, Dec 2021.

[23.791] 3GPP TR 23.791, “Study of Enablers for Network Automation for 5G," Release 16, v16.2.0, June 2019.

[28.533] 3GPP TS 28.533, “Management and Orchestration; Architecture framework," Release 17, v17.1.0, Dec 2021.

[33.501] 3GPP TS 33.501 “Security architecture and procedures for 5G system,” Release 17, v17.4.2, Jan 2022.

[35.206] 33GPP TS 35.206 “3rd Generation Partnership Project; Technical Specification Group Services and System Aspects; 3G Security; Specification of the MILENAGE Algorithm Set: An example algorithm set for the 3GPP authentication and key generation functions f1, f1*, f2, f3, f4, f5 and f5*; Document 2: Algorithm Specification," Release 16, v16.0.0, July 2020.

[35.231] 3GPP TS 35.231 “Specification of the TUAK algorithm set: A second example algorithm set for the 3GPP authentication and key generation functions f1, f1*, f2, f3, f4, f5 and f5*; Document 1: Algorithm specification,” Release 16, v16.0.0, July 2020.

[38.331] 3GPP TS 38.331 “NR; Radio Resource Control (RRC); Protocol specification,” Release 16, v16.7.0, Dec 2021.

[5GA19] 5G-ACIA Whitepaper “5G for Connected Industries and Automation,” Second edition, 2019.

[5GP] The 5G Infrastructure Public Private Partnership (5G-PPP), https://5g-ppp.eu

[5GP20] 5G-PPP, “Empowering vertical industries through 5G networks – Current status and future trends,” 5G-PPP and 5GIA, 2020, available online at: http://doi.org/10.5281/zenodo.3698113

[5GP21] 5G-PPP, "Service performance measurement methods over 5G experimental networks", May 2021, available online at:

https://5g-ppp.eu/wp-content/uploads/2021/06/Service-performance-measurement-methods-over-5G-experimental-networks_08052021-Final.pdf

[5GP21] Gavras, Anastasius et al., “5G PPP Architecture Working Group - View on 5G Architecture, Version 4.0”, Oct. 2021, available online at: 10.5281/ZENODO.5155657.

[5GPa] 5G-PPP, “PHASE 3.6: 5G INNOVATIONS AND BEYOND 5G”, available online at: https://5g-ppp.eu/5g-ppp-phase-3-6-projects/

[AAFO] AI Agents for OSM’s documentation, available online at: https://ai-agents-for-osm.readthedocs.io/en/latest.

[ADE] ADEME, “Quantifying the impact of an emission reduction action on GHGs”, version 2

[AGB+20] K. Antevski, M. Groshev, G. Baldoni, and et al., “DLT federation for Edge robotics,” IEEE Conference on Network Function Virtualization and Software Defined Networks (NFV-SDN), IEEE, 2020, pp. 71–76.

[ALL+19] Aguado, A., Lopez, V., Lopez, D., et al., “The engineering of software-defined quantum key distribution networks”, IEEE Communications Magazine, 57(7), 20-26, 2019.

[AHT+20] R.A.F.Alvarengaa, S.Huysvelda, S.E.Taelmana, et. Al., “A framework for using the handprint concept in attributional life cycle (sustainability) assessment”, Journal of Cleaner Production, Volume 265, 2020, available online at: https://doi.org/10.1016/j.jclepro.2020.121743

[AMB20] N. Aydın, İ. Muter, and Ş. İ. Birbil, “Multi-objective temporal bin packing problem: An application in cloud computing,” Computers & Operations Research, vol. 121, 2020.

[AND20] Andrae, A.S. “Hypotheses for primary energy use, electricity use and CO2 emissions of global computing and its shares of the total between 2020 and 2030”, WSEAS Transactions on Power Systems, 2020.

[ANS21] ANSSI v1.0, "Managing Cybersecurity for Industrial Control Systems," June 2021.

[Arc20] Arcep, “Networks and the Environment”, Sep 2020, available online at: https://en.arcep.fr/news/press-releases/view/n/networks-and-the-environment.html

[Arc20a] Archyde, “Paris gives itself four months to decide”, Nov 2020, available online at: https://www.archyde.com/paris-gives-itself-four-months-to-decide/

[ARCEP19] French regulator ARCEP 2019 study on the environmental impact of the digital technologies, available online at: https://www.arcep.fr/uploads/tx_gspublication/reseaux-du-futur-empreinte-carbone-numerique-juillet2019.pdf

[ALB+04] Avizienis, Algirdas, J-C. Laprie, Brian Randell, et al., "Basic concepts and taxonomy of dependable and secure computing." IEEE transactions on dependable and secure computing 1, no. 1, 11-33, 2004.

[AYM+20] P. Ahokangas, S. Yrjölä, M. Matinmikko-Blue, et al., ”Antecedents of Future 6G Mobile Ecosystems,” 2nd 6G Wireless Summit (6G SUMMIT), Levi, Finland, 2020.

[BDG+14] P.Bosshart, D.Daly, G.Gibb, et al., "P4: Programming protocol-independent packet processors." ACM SIGCOMM Computer Communication Review 44, no. 3, pp. 87-95, 2014.

[BEL+18] Belkhir L., Ahmed E. “Assessing ICT global emissions footprint: trends to 2040 & recommendations”, J. Clean. Prod. 177:448–463, 2018.

[BVM+20] Bergmark, Pernilla, Vlad C. Coroamă, Mattias Höjer, et al., "A methodology for assessing the environmental effects induced by ict services: Part ii: Multiple services and companies." In Proceedings of the 7th International Conference on ICT for Sustainability, pp. 46-55. 2020.

[BFC21] R. Bassoli, F. H.P. Fitzek, and E. Calvanese Strinati, "Why Do We Need 6G?" ITU Journal on Future and Evolving Technologies, vol. 2 no. 6, Sep. 2021.

[BGL+] Boubendir, A., Guillemin, F., Le Toquin, C., et al., "Federation of cross-domain edge resources: a brokering architecture for network slicing", In 4th IEEE Conference on Network Softwarization and Workshops (NetSoft) (pp. 415-423). IEEE, 2018.

[BGS+20] R. Bassoli, F. Granelli, C. Sacchi, et al., "CubeSat-Based 5G Cloud Radio Access Networks: A Novel Paradigm for On-Demand Anytime/Anywhere Connectivity," IEEE Vehicular Technology Magazine, vol. 15, no. 2, pp. 39-47, June 2020, available online at: 10.1109/MVT.2020.2979056.

[BGT+18] A. Boubendir, F. Guillemin, C. Le Toquin, et. Al., "Federation of cross-domain edge resources: a brokering architecture for network slicing," 4th IEEE Conference on Network Softwarization and Workshops (NetSoft), pp. 415-423, 2018.

[BSH+20] J. Bieser, B. Salieri, R. Hischier, et al., "Next generation mobile networks Problem or opportunity for climate protection?" University of Zurich, Empa, 2020.

[BME+20] J. Baranda; J. Mangues-Bafalluy; Engin Zeydan, et al., "On the Integration of AI/ML-based scaling operations in the 5Growth platform," IEEE Conference on Network Function Virtualization and Software Defined Networks (NFV-SDN), pp. 105-109, 2020, available online at: 10.1109/NFV-SDN50289.2020.9289863

[BMWi19] Federal Ministry for Economic Affairs and Energy (BMWi), "Shaping Digital Ecosystems Globally - 2030 Vision for Industrie 4.0," Whitepaper, 2019, available online at:

https://www.plattform-i40.de/IP/Redaktion/EN/Downloads/Publikation/Vision-2030-for-Industrie-4.0.html

[BMZ+20] J. Baranda, J. Mangues-Bafalluy, E. Zeydan, et al., "On the Integration of AI/ML-based scaling operations in the 5Growth platform," IEEE Conference on Network Function Virtualization and Software Defined Networks (NFV-SDN), pp. 105-109, 2020, available online at: 10.1109/NFV-SDN50289.2020.9289863

[BFY+21] Bomin Mao, Fengxiao Tang, Yuichi Kawamoto, et al., "AI based Service Management for 6G Green Communications," available online at: https://arxiv.org/pdf/2101.01588.pdf, 2021

[BR18] R. Bifulco and G. Rétvári,"A survey on the programmable data plane: Abstractions, architectures, and open problems." IEEE 19th International Conference on High Performance Switching and Routing (HPSR). IEEE, 2018.

[BT20] C. Benzaid and T. Taleb, "AI-driven zero touch network and service management in 5Gand beyond: Challenges and research directions," IEEE Network 34, no. 2, pp. 186–194, 2020.

[BT20a] C. Benzaid and T. Taleb, "ZSM Security: Threat Surface and Best Practices", IEEE Network, vol. 34, no. 3, pp. 124–133, May 2020, available online at: 10.1109/MNET.001.1900273

[CCC] "Confidential Computing: The Next Frontier in Data Security," White Paper, Confidential Computing Consortium, https://confidentialcomputing.io/white-papers-reports/.

[CCC20] "Confidential Computing: Hardware-Based Trusted Execution for Applications and Data", available online at:

https://confidentialcomputing.io/wpcontent/uploads/sites/85/2020/06/ConfidentialComputing_OSSNA2020.pdf

[CHAN18] P. Chang and G. Miao, "Resource Provision for Energy-Efficient Mobile Edge Computing Systems," in IEEE Global Communications Conference (GLOBECOM), pp. 1–6, 2018.

[CDB21] L. Chiaraviglio, C. Di Paolo, and N. Blefari Melazzi, "5G Network Planning under Service and EMF Constraints: Formulation and Solutions," IEEE Transactions on Mobile Computing, pp. 1–1, 2021.

[CBM+20] Coroamă, Vlad C., Pernilla Bergmark, Mattias Höjer, et al., "A methodology for assessing the environmental effects induced by ict services: Part i: Single services." In Proceedings of the 7th International Conference on ICT for Sustainability, pp. 36-45. 2020.

[CPT21] CEPT, ECC Report Draft: "Receiver selectivity performance of satellite Earth stations in the band 3800-4200 MHz," 2021, available online at: https://www.cept.org/Documents/se-40/66427/se4021035_annex6_rev4_wd_eccreport_c_band_performance_thursday_final

[CRA+21] D. Canastro, R. Rocha, M. Antunes, et al., "Root Cause Analysis in 5G/6G Networks," 8th International Conference on Future Internet of Things and Cloud (FiCloud), IEEE, pp. 217-224, 2021.

[DDA+15] A. De Domenico, L.-F. Diez, R. Aguero, et al., "EMF-Aware Cell Selection in Heterogeneous Cellular Networks," IEEE Communications Letters, vol. 19, no. 2, pp. 271–274, 2015.

[DLH19] Y. Dang, Q. Lin, and P. Huang, "Aiops: real-world challenges and research innovations," 41st International Conference on Software Engineering: Companion Proceedings (ICSE-Companion), IEEE, pp. 4-5, 2019.

[DPS20] E. Dahlman, S. Parkvall, and J. Sköld, "5G NR: The Next Generation Wireless Access Technology", 2nd edition, Academic Press, 2020, ISBN-13: 978-0128223208.

[DR14] C. Dwork and A. Roth, "The algorithmic foundations of differential privacy," Foundation and Trends in Theoretical Computer Science, vol. 9, no. 3–4, pp. 211–407, 2014, available online at: 10.1561/0400000042.

[ETS17a] ETSI EN 303 146-4 V1.1.2, "Radio Virtual Machine (RVM)" developed by the European Telecommunications Standards Institute, 2017.

[ERI20] Building trustworthiness into future mobile networks, available online at: https://www.ericsson.com/en/reports-and-papers/white-papers/building-trustworthiness-into-future-mobile-networks

[ETS14] ETSI ES 203 199, "Methodology for environmental Life Cycle Assessment (LCA) of Information and Communication Technology (ICT) goods, networks and services", 2014.

[ETS17] ETSI GS NFV-SEC 013 V3.1.1 “Network Functions Virtualisation (NFV) Release 3; Security; Security Management and Monitoring specification," February 2017.

[ETS19] ETSI GS ZSM 001 V1.1.1, “Zero-touch network and Service Management (ZSM); Requirements based on documented scenarios," Oct. 2019.

[ETS19a] ETSI GS ZSM 002 V1.1.1 “Zero-touch network and Service Management (ZSM); Reference Architecture," August 2019.

[ETS20] ETSI GR PDL 003 v1.1.1, “Permissioned Distributed Ledger (PDL); Application Scenarios,” Dec. 2020.

[ETS20a] ETSI TR 103 477 V1.2.1, “eHEALTH; Standardization use cases for eHealth,” Aug. 2020.

[ETS21] ETSI GS NFV-SEC 024 V0.0.6, “Network Functions Virtualisation (NFV); Security; Security Management Release 4,"DRAFT, April 2021.

[ETS21a] ETSI GS ZSM 009-1 V1.1.1, “Zero-touch network and Service Management (ZSM); Closed-Loop Automation; Part 1: Enablers," June 2021.

[ETS21b] ETSI GR ZSM 010 V1.1.1, “Zero-touch network and Service Management (ZSM); General Security Aspects,” July 2021.

[ETS21c] ETSI GR NFV-IFA 039 V0.0.20 “Network Functions Virtualisation (NFV); Release 4 Architectural Framework; Report on Service Based Architecture (SBA) design," DRAFT, November 2021.

[FBF+17] A. Francescon, G. Baggio, R. Fedrizzi, et al., “X–MANO: Cross–domain management and orchestration of network services,” IEEE Conference on Network Softwarization (NetSoft), IEEE, pp. 1–5, 2017.

[FCC19] FCC, Spectrum Horizons: A Rule by the Federal Communications Commission on 06/04/2019, available online at:

https://www.federalregister.gov/documents/2019/06/04/2019-10925/spectrum-horizons

[FGS20] F. Fitzek, F. Granelli, and P. Seeling, “Computing in Communication Networks – From Theory to Practice,” 1st edition, Elsevier, 2020, ISBN: 9780128204887.

[FLS+21] F. Fitzek, S.C. Li, S. Speidel, et al., “Tactile Internet with Human-in-the-Loop", Academic Press, 2021.

[FBW+20] C. Freitag, M. Berners-Lee, K. Widdicks, et al., available online at: https://www.ericsson.com/en/blog/2020/2/climate-impact-of-digital-technology

[FRE+21] C. Freitag, M. Berners-Lee, K. Widdicks, B. Knowles, G. S. Blair and A. Friday, The real climate and transformative impact of ICT: A critique of estimates, trends, and regulations, Patterns (N Y). 2021 Sep 10;2(9):100340. doi: 10.1016/j.patter.2021.100340.

[FRE21] P. Frenger and K. W. Helmersson, "Massive MIMO Muting using Dual-polarized and Array-size Invariant Beamforming," IEEE 93rd Vehicular Technology Conference (VTC2021-Spring), pp. 1-6, 2021, available online at: 10.1109/VTC2021-Spring51267.2021.9448654.

[GAL21] J. Galan-Jimenez and L. Chiaraviglio, “Measuring the impact of ICNIRP vs. stricter-than-ICNIRP exposure limits on QoS and EMF from cellular networks,” Computer Networks, vol. 187, p.107824, 2021, available online at: https://www.sciencedirect.com/science/article/pii/S1389128621000128

[GFR20] J. García-Morales, G. Femenias and F. Riera-Palou, "Energy-Efficient Access-Point Sleep-Mode Techniques for Cell-Free mmWave Massive MIMO Networks with Non-Uniform Spatial Traffic Density," in IEEE Access, vol. 8, pp. 137587-137605, 2020, available online at: 10.1109/ACCESS.2020.3012199.

[GSS+19] Gati, Azeddine, Fatma Ezzahra Salem, Ana Maria Galindo Serrano, et al., "Key technologies to accelerate the ICT Green evolution--An operator's point of view." arXiv preprint arXiv:1903.09627, 2019.

[GRA18] F. Granelli and R. Bassoli, "Autonomic Mobile Virtual Network Operators for Future Generation Networks," IEEE Network, vol. 32, no. 5, pp. 76-84, September/October 2018, available online at: 10.1109/MNET.2018.1700455.

[GRA18a] F. Granelli and R. Bassoli, "Towards Autonomic Mobile Network Operators," IEEE 7th International Conference on Cloud Networking (CloudNet), pp. 1-4, 2018, available online at: 10.1109/CloudNet.2018.8549552.

[GBD+18] Z. Guan, L. Bertizzolo, E. Demirors, and T. Melodia, "WNOS: An Optimization-based Wireless Network Operating System", Proceedings of the Eighteenth ACM International Symposium on Mobile Ad Hoc Networking and Computing, June 2018.

[GCZ+20] F. Granelli, C. Costa, J. Zhang, et al., "Design of an On-Demand Agile 5G Multi-Access Edge Computing Platform Using Aerial Vehicles," in IEEE Communications Standards Magazine, vol. 4, no. 4, pp. 34-41, December 2020, available online at: 10.1109/MCOMSTD.001.2000016.

[GDPR] GDPR Key Issues: Privacy by Design, Available at: https://gdpr-info.eu/issues/privacy-by-design/

[GHGP11] World Resources Institute and World Business Council for Sustainable Development, Corporate Value Chain (Scope 3) Accounting and Reporting Standard, September 2011 ISBN 978-1-56973-772-9

[GKM18] A. Grange, I. Kacem, and S. Martin, “Algorithms for the bin packing problem with overlapping items,” Computers & Industrial Engineering, vol. 115, pp. 331-341, 2018.

[GoogleAI] https://ai.googleblog.com/

[GSM18] Global System for Mobile Communications (GSMA), “Mobile Industry Impact Report: Sustainable Development Goals,” 2018.

[GSM21] GSMA, Industry pathways to net zero: mobile and digital technology in support of industry decarbonisation, 2021.

[GSM19] Energy Efficiency: An Overview, May 2019, available online at: https://www.gsma.com/futurenetworks/wiki/energy-efficiency-2/ Accessed: 2022-02-22

[GSM21] GSMA V1.0, “Quantum Computing, Networking and Security,” March 2021, available online at: https://www.gsma.com/newsroom/wp-content/uploads//IG-11-Quantum-Computing-Networking-and-Security.pdf.

[GSM] GSMA, “The Enablement Effect, the impact of mobile communications technologies on carbon emission reduction”, available online at: https://www.gsma.com/betterfuture/wp-content/uploads/2019/12/GSMA_Enablement_Effect.pdf

[GSS14] Goodfellow, Ian J., Jonathon Shlens, and Christian Szegedy. "Explaining and harnessing adversarial examples." arXiv preprint arXiv:1412.6572 (2014).

[HEX21-D12] Hexa-X, “Deliverable D1.2: Expanded 6G vision, use cases and societal values – including aspects of sustainability, security and spectrum”, April 2021.

[HEX21-D21] Hexa-X, “Deliverable D2.1: Towards Tbps Communications in 6G: Use cases and Gap Analysis”, June 2021.

[HEX21-D22] Hexa-X, “Deliverable D2.2: Initial radio models and analysis towards ultra-high data rate links in 6G”, Jan. 2022.

[HEX21-D31] Hexa-X, “Deliverable D3.1: Localisation and sensing use cases and gap analysis”, Jan. 2022.

[HEX21-D41] Hexa-X, “Deliverable D4.1: AI-driven communication & computation co-design: Gap analysis and blueprint”, August 2021.

[HEX21-D51] Hexa-X, “Deliverable D5.1: Initial 6G Architectural Components and Enablers”, December 2021.

[HEX21-D71] Hexa-X, “Deliverable D7.1: Gap analysis and technical work plan for special-purpose functionality”, July 2021.

[HJS17] H. Hawilo, M. Jammal and A. Shami, "Orchestrating network function virtualization platform: Migration or re-instantiation?" IEEE 6th International Conference on Cloud Networking (CloudNet), pp. 1-6, 2017, available online at:: 10.1109/CloudNet.2017.8071528.

[HPE+21] J. L. Hevia, G. Peterssen, C. Ebert, et al., "Quantum Computing," IEEE Software, vol. 38, no. 5, pp. 7-15, September-October 2021.

[HUA21] O. Haliloglu, E. Ustundag Soykan, A. Alabbasi, “Privacy Preserving Federated RSRP Estimation for Future Mobile Networks,” IEEE Globecom workshop (GC Wkshps) on 6G Oriented Trustworthiness, pp. 1-6, 2021.

[HSM] HSM security in 5G core networks, available online at: https://www.ericsson.com/en/core-network/5g-core/hsm-security

[Hue08] M. C. Huebscher and J. A. McCann, “A survey of autonomic computing—degrees, models, and applications”, ACM Comput. Surv., vol. 40, no. 3, pp. 1–28, August 2008, available online at: 10.1145/1380584.1380585.

[IBM05] IBM whitepaper, "An architectural blueprint for autonomic computing," Third edition, June 2005.

[IBQ+19] Interdonato, G., Björnson, E., Quoc Ngo, et al. “Ubiquitous cell-free Massive MIMO communications.” J Wireless Com Network, 197, 2019.

[IQQ] “ID Quantique Quantis QRNG Chip”, available online at: https://www.idquantique.com/random-number-generation/products/quantis-qrng-chip/

[IEEE19] IEEE Standard for a Precision Clock Synchronization Protocol for Networked Measurement and Control Systems, available online at: https://standards.ieee.org/ieee/1588/4355/

[IFL19] G. Interdonato, P. Frenger and E. G. Larsson, "Scalability Aspects of Cell-Free Massive MIMO," ICC IEEE International Conference on Communications (ICC), 2019, pp. 1-6, available online at: 10.1109/ICC.2019.8761828.

[INS21] INSPIRE-5Gplus, “D3.2: Security drivers and associated software-defined models”, Version 1.3, November 2021.

[ISO20] ISO, “Automation systems and integration — Digital Twin framework for manufacturing. Overview and general principles,” Oct. 2020.

[ISO06a] ISO 14040:2006, Environmental management — Life cycle assessment — Principles and framework, 2006.

[ISO06b] ISO 14044:2006, Environmental management — Life cycle assessment — Requirements and guidelines, 2006.

[ISO15] ISO/IEC 11889-1:2015 Information technology — Trusted Platform Module — Part 1: Overview

[ISO18] ISO 14064-1:2018, Greenhouse gases — Part 1: Specification with guidance at the organization level for quantification and reporting of greenhouse gas emissions and removals, 2018.

[ITU12] ITU L.1420 Methodology for energy consumption and greenhouse gas emissions impact assessment of information and communication technologies in organizations, 2012.

[ITU14] ITU L.1410 Methodology for environmental life cycle assessments of information and communication technology goods, networks and services, 2014.

[ITU18] ITU L.1450 Methodologies for the assessment of the environmental impact of the information and communication technology sector, 2018.

[ITU20] ITU L.1470 Greenhouse gas emissions trajectories for the information and communication technology sector compatible with the UNFCCC Paris Agreement, 2020.

[JAS02] G. Joya, M. A., Atencia, and F. Sandoval, “Hopfield neural networks for optimization: study of the different dynamics,” Neurocomputing vol. 43, no. 1-4, pp. 219-237, 2002.

[JKM+21] Jegorova, Marija, Chaitanya Kaul, Charlie Mayor, et al., "Survey: Leakage and privacy at inference time." arXiv preprint arXiv:2107.01614, 2021.

[VSP+22] Valero, José María Jorquera, Pedro Miguel Sánchez Sánchez, Manuel Gil Pérez, et al., "Toward pre-standardization of reputation-based trust models beyond 5G." Computer Standards & Interfaces 81: 103596, 2022.

[KOB21] F. Karakoç, M. Önen, and Z. Bilgin “Secure Aggregation Against Malicious Users,” 26th ACM Symposium on Access Control Models and Technologies, pp. 115-124, 2021.

[LLF+21] Z. Lv, R. Lou, H. Feng, et al., "Novel machine learning for big data analytics in intelligent support information management systems," ACM Transactions on Management Information System (TMIS), vol. 13, no. 1, pp. 1-21, 2021.

[LAG19] A. Laghrissi and T. Taleb, "A Survey on the Placement of Virtual Resources and Virtual Network Functions," in IEEE Communications Surveys & Tutorials, vol. 21, no. 2, pp. 1409-1434, Second quarter 2019, available online at: 10.1109/COMST.2018.2884835.

[LIM21] B. Lim, S. Zohren, “Time-series forecasting with deep learning: a survey,” Philosophical Transactions of the Royal Society A, 379, no. 2194, 2021. Available online at:

https://www.oxford-man.ox.ac.uk/wp-content/uploads/2020/11/Time-Series-Forecasting-With-Deep-Learning_-A-Survey.pdf

[MAA+20] M. Matinmikko-Blue, S. Aalto, M. Asghar, et al., "White Paper on 6G Drivers and the UN SDGs,” White paper, 6G Research Visions 2, University of Oulu, Finland, 2020, available online at: http://urn.fi/urn:isbn:9789526226699

[MBM+20] N. H. Mahmood, S. Böcker, A. Munari, et al., “White paper on critical and massive machine type communication towards 6G,” arXiv preprint arXiv:2004.14146, 2020.

[MAL18] Malmodin J., Lundén D. “The energy and carbon footprint of the global ICT and E&M sectors 2010–2015”. Sustainability, 10:3027, 2018.

[MAL18a] Malmodin J., Lundén D. Report from the KTH Centre for Sustainable Communications Stockholm; The Electricity Consumption and Operational Carbon Emissions of ICT Network Operators 2010-2015”, Technical report, 2018.

[MSL+20] E. Masanet, A. Shehabi, N. Lei, et al., “Recalibrating global data center energy-use estimates”, Science Vol 367 Issue 6481, 2020.

[MAS+19] V. Martin, A. Aguado, P. Salas, et al., "The Madrid Quantum Network: A Quantum-Classical Integrated Infrastructure," in OSA Advanced Photonics Congress, OSA Technical Digest (Optica Publishing Group, 2019), paper QtW3E.5, 2019.

[MBB+20] A. Molina Zarca, M. Bagaa, J. Bernal Bernabe, et al., “Semantic-Aware Security Orchestration in SDN/NFV-Enabled IoT Systems”, Sensors, vol. 20, no. 13, p. 3622, June 2020, available online at: 10.3390/s20133622.

[ETS22] Draft ETSI GS MEC 040, " Multi-access Edge Computing (MEC); Federation enablement APIs", v3.0.8, Feb. 2022, available online at: https://docbox.etsi.org/ISG/MEC/Open/MEC040%20FederationAPI%20drafts/gs_mec040federationapiv308_Early%20draft.pdf

[MDD+21] M. Merluzzi, N. di Pietro, P. Di Lorenzo, et al., "Discontinuous Computation Offloading for Energy-Efficient Mobile Edge Computing," in IEEE Transactions on Green Communications and Networking, available online at: 10.1109/TGCN.2021.3125543., 2021

[MRG11] Meddour, Djamal-Eddine, Tinku Rasheed, and Yvon Gourhant. "On the role of infrastructure sharing for mobile network operators in emerging markets." Computer networks 55, no. 7: 1576-1591, 2011.

[NIS18] NIST v1.1, "Framework for Improving Critical Infrastructure Cybersecurity,” April 2018, available online at: https://doi.org/10.6028/NIST.CSWP.04162018.

[OFC21] OFCOM, Spectrum Access: EHF licences, 2021, available online at: https://www.ofcom.org.uk/manage-your-licence/radiocommunication-licences/spectrum-access-ehf

[PGB+21] Palamà, I., Gringoli, F., Bianchi, et al., “IMSI Catchers in the wild: A real world 4G/5G assessment”, Computer Networks, 194, 108137, 2021.

[MUS17] M.-S. Mushtaq and A. Mellouk, “Quality of Experience Paradigm in Multimedia Systems”. Elsevier 2017, available online at: https://doi.org/10.1016/C2015-0-05972-X

[RAD78] R. L. Rivest, L. Adleman, and M. L. Dertouzos, “On data banks and privacy homomorphisms.” Foundations of Secure Computation, vol. 4, no. 11, pp. 169-180, 1978.

[RBL+17] B. Rodrigues, T. Bocek, A. Lareida, et al., “A blockchain-based architecture for collaborative DDoS mitigation with smart contracts,” IFIP International Conference on Autonomous Infrastructure, Management and Security, Springer, Cham, pp. 16-29, 2017.

[RIT21] Ritchie, H., Roser, M. "CO_2 and Greenhouse Gas Emissions", OurWorldInData.org, 2020, available online at: https://ourworldindata.org/co2-and-other-greenhouse-gas-emissions

[RLM18] R. Roman, J. Lopez, and M. Mambo, "Mobile Edge Computing, Fog et al.: A Survey and Analysis of Security Threats and Challenges", *Future Generation Computer Systems*, vol. 78, pp.680–698, Jan. 2018, available online at: 10.1016/j.future.2016.11.009.

[ROU21] Roussilhe, G. Que peut le numérique pour la transition écologique? (not translated), 2021.

[RR20] ITU, Radio Regulations, 2020.

[RWT13] Ruefle, Robin, Ken van Wyk, et Lana Tosic, "New Zealand Security Incident Management Guide for Computer Security Incident Response Teams (CSIRTs)", May 2013.

[SCA+19] F. E. Salem, T. Chahed, E. Altman, et al., "Optimal Policies of Advanced Sleep Modes for Energy-Efficient 5G networks," IEEE 18th International Symposium on Network Computing and Applications (NCA), 2019, pp. 1-7, 2019, available online at: 10.1109/NCA.2019.8935062.

[SSS11] M. Sindelar, R. K. Sitaraman, and P. Shenoy, "Sharing-Aware Algorithms for Virtual Machine Colocation," Proceedings of the twenty-third annual ACM symposium on Parallelism in algorithms and architectures, pp. 367-378, 2011.

[SZF+18] T. Szigeti, D. Zacks, M. Falkner, et al., "Cisco Digital Network Architecture: Intent-based Networking for the Enterprise," Cisco Press, 2018.

[SZI21] P. Szilágyi, "I2bn: Intelligent intent-based networks," Journal of ICT Standardization, pp. 159-200, 2021.

[TBH+19] E. Tabassi, K. Burns, M. Hadjimichael, et al., "A Taxonomy and Terminology of Adversarial Machine Learning," NIST, Oct. 2019.

[TCP91] G. A. Tagliarini, J. F., Christ, and E. W. Page, "Optimization using neural networks," IEEE transactions on computers, vol. 40, no. 12, pp. 1347-1358, 1991.

[TCD+14] Tesanovic, Milos, Emmanuelle Conil, Antonio De Domenico, et al., "The LEXNET project: Wireless networks and EMF: Paving the way for low-EMF networks of the future." IEEE Vehicular Technology Magazine 9, no. 2, pp. 20-28, 2014.

[TMF] TM Forum, Autonomous Networks Industry Standards, v1.1.0, https://www.tmforum.org/resources/how-to-guide/ig1230b-autonomous-networks-industry-standards-v1-1-0.

[UZK+20] R.B. Uriarte, H. Zhou, K. Kritikos, et al., "Distributed service-level agreement management with smart contracts and blockchain," Concurrency and Computation: Practice and Experience, vol. 33, no. 14, p. e5800, 2020.

[VAS16] A. Vassilev and R. Staples, "Entropy as a Service: Unlocking Cryptography's Full Potential," in Computer, vol. 49, no. 9, pp. 98-102, Sept. 2016, available online at: 10.1109/MC.2016.275.

[WZY+18] S. Wang, X. Zhang, Z. Yan, et al., "Cooperative Edge Computing with Sleep Control under Non-uniform Traffic in Mobile Edge Networks," IEEE Internet Things J., Oct. 2018.

[YAO82] A. C. Yao, “Protocols for secure computations.” 23rd annual symposium on foundations of computer science (sfcs 1982), IEEE, pp. 160-164, 1982.

[YYY+19] Q. Yang, Y. Liu, Y. Cheng, et al., “Federated learning. Synthesis Lectures on Artificial Intelligence and Machine Learning,” vol. 13, no. 3, pp. 1-207, 2019.

[ZSV+21] V. Ziegler, P. Schneider, H. Viswanathan, et al., “Security and Trust in the 6G Era”, IEEE Access, vol. 9, pp. 142314–142327, 2021, available online at: 10.1109/ACCESS.2021.3120143.

Annex A: Terminologies

Term	Acronym	Term description	Reference
Angle-of-arrival (also direction-of-arrival)	AoA / DoA	Measurement of Angle (in azimuth and/or elevation) of a signal incoming to an array from a certain direction, measured in the coordinate system of that array.	D3.1
Abstraction Model	N/A	Process of focusing on the important characteristics and behaviour of a concept and realizing this as a set of one or more elements in an information or data model.	https://www.etsi.org/deliver/etsi_gr/ENI/001_099/004/02.01.01_60/gr_ENI004v020101p.pdf
Accuracy, precision, and resolution of a measurement	N/A	Different definitions exist in the literature. Accuracy can refer to the statistical bias (difference between the mean of the measurements and the true value), but also to the percentile performance (e.g., a positioning system with 95% accuracy of 5 meters indicates that 95% of the errors are within 5 meters). Precision refers to the spread of the measurements around the mean and is thus related to the (co-)variance of the measurements. Finally, resolution refers to the ability of the measurement system to distinguish nearby signal sources (e.g., a radar with 1 GHz bandwidth has a resolution of 0.15 m so that two objects with a distance exceeding 15 cm can be distinguished).	D3.1
Active learning	N/A	The selecting of important samples from a large unlabelled dataset for labelling (by querying an oracle) to accelerate model training with a labelling budget.	D4.1
AI agent	N/A	Network entity carrying one or more (trained) ML models; it can be instantiated across the network, i.e., at network infrastructure or device side.	D4.1
AI agent availability	N/A	Availability (or readiness) of an AI agent to accept inferencing requests and address them with high accuracy.	D4.1
AI agent density	N/A	Density of devices with AI/ML components considering specific traffic patterns during data sharing.	D4.1

AI agent reliability	N/A	Capability of an AI agent to accept inferencing requests and provide high accuracy output in a timely manner (within a deadline set by the requesting application).	D4.1
AI deployment flexibility	N/A	Flexibility to deploy same system in multiple scenarios without many modifications to the AI models. Goes hand in hand with generalisability	D4.1
AI function	N/A	Functional entity in which an AI agent carrying a (trained) ML model can be instantiated.	D4.1
AI orchestration function	N/A	AI function functionality for AI agent discovery and selection.	D4.1
AI policy enforcer	N/A	AI function functionality to implement a recommended policy.	D4.1
AI success monitoring function	N/A	AI function functionality performing inferencing accuracy, communication efficiency, and security monitoring.	D4.1
AI system transparency	N/A	Property that enables interpretability and, thus, the explanation of the decision-making process of AI system.	D4.1
AI Trustworthiness	N/A	AI-based models should perform optimally as intended by design without any unauthorised manipulation.	D4.1
AI/ML fairness	N/A	Ability of the AI/ML agent to perform a decision free from discrimination and bias.	D4.1
AI/ML flexibility	N/A	Ability of the ML model to adapt to different conditions/environments in a timely fashion.	D4.1
AI-enabled network energy efficiency	N/A	Training/inference energy optimisation in edge/IoT ecosystem.	D4.1
Analog-to-Digital Converter	ADC	An electronic integrated circuit used to translate analog electrical signals to digital or binary form consisting of 1s and 0s	D2.1
Analytics Logical Function	AnLF	Function that performs inference, derives analytics information, and exposes analytics service.	D4.1
Angle-of-departure (also direction-of-departure)	AoD / DoD	Measurement Angle (in azimuth and/or elevation) of a signal outgoing from an array towards a corresponding direction, measured in the coordinate system of that array.	D3.1

Antenna gain	N/A	The ratio, usually expressed in decibels, of the power required at the input of a loss-free isotropic reference antenna to the power supplied to the input of the given antenna to produce, in a given direction, the same field strength or the same power flux-density at the same distance. When not specified otherwise, the gain refers to the direction of maximum radiation.	D2.1
Application (software/program)	N/A	Software that is specific to the solution of a problem usually submitted by an end user. For clarification prefixes could be used, e.g., end user application, network application. Note: Application (software/ program) and service is often used as synonym. In our context application is used for software with user interface (UI).	D3.1
Architecture (Rel. to Software Systems)	N/A	Defines the high-level structure and organization of a software-based system. This includes the objects, their properties and methods, and relationships between objects.	https://www.etsi.org/deliver/etsi_gr/ENI/001_099/004/02.01.01_60/gr_ENI004v020101p.pdf
Artificial Intelligence	AI	The science and engineering of making computer behave in ways that, until recently, we thought that required human intelligence.	D4.1
Availability	N/A	Availability is indicated by the percentage of time during which all required QoS parameters are satisfied (correct operation). The required QoS parameters and their target ranges are service and use case specific.	D7.1
Basic transmission loss (of a radio link) $\boldsymbol{L_b}$	N/A	The ratio, usually expressed in decibels, for a radio link between the power radiated by the transmitting antenna and the power that would be available at a conjugately matched receiver antenna input if the antennas were replaced by isotropic antennas with the same polarisation as the real antennas, including the attenuation effects on the propagation path, but with the effects of obstacles close to the antennas being disregarded. $L_b = L_{bf} + L_m$ [dB] where L_m is the loss relative to free space. The loss relative to free space, L_m, may be divided into losses of	D2.1

		different types, such as: absorption loss (ionospheric or atmospheric gases, precipitation, clouds, etc.); diffraction loss as for ground waves; effective reflection or scattering loss as in the ionospheric case including the results of any focusing or defocusing due to curvature of a reflecting layer; polarisation coupling loss; this can arise from any polarisation mismatch between the antennas for the particular ray path considered; aperture-to-medium coupling loss or antenna gain degradation, which may be due to the presence of substantial scatter phenomena on the path; beam spreading loss; effect of wave interference between the direct ray and rays reflected from the ground, other obstacles or atmospheric layers; clutter loss; building entry loss.	
Beamforming	N/A	A signal processing technique used in combination with antenna arrays to achieve a defined directional antenna characteristic enabling directional signal transmission or reception. This is achieved by individual weighting of the signals of each antenna element in such a way that signals at particular angular directions experience constructive interference while others experience destructive interference.	D2.1
Bistatic sensing	N/A	Sensing (see Sensing), whereby the transmitting and receiving nodes are not co-located.	D3.1
Campus network	N/A	A campus network is a network made up of an interconnection of Local Area Network (LANs) within a limited geographical area.	D5.1
Channel hardening	N/A	A fading channel behaves as if it was a non-fading channel	D2.1
Classification problem	N/A	Problem that requires the algorithm to associate its input to one or several discrete target values, often called labels.	D4.1
Closed loop control	N/A	Self-regulating mechanism in which outputs of a system are provided to a system that compares the current state to a desired state (or set of states); the comparison is then used to adjust the behaviour of the system	https://www.etsi.org/deliver/etsi_gr/ENI/001_099/004/02.01.01_60/gr_ENI004v020101p.pdf

Cloud Computing	N/A	A model that offers compute resources like CPU, network, and disk capabilities on-demand over the internet. Cloud computing gives users the ability to access and use computing power in a remote physical location	https://github.com/cncf/glossary/blob/main/content/en/cloud_computing.md
Cloud Network	N/A	A computer network within or make part of a cloud computing infrastructure.	D6.1
Cloud-native	N/A	It is an architecture model to build and run software applications by exploiting the scalability, and resilience of cloud computing.	D6.1
Coherent Processing Interval	CPI	Duration in which a signal is received and during which the geometric parameters remain constant.	D3.1
Communications Services Provider	CSP	Company or organisation, making use of an electronics communications network or part thereof to provide a service or services on a commercial basis to third parties.	https://www.etsi.org/deliver/etsi_gs/NFV/001_099/003/01.03.01_60/gs_NFV003v010301p.pdf
Compute-as-a-Service	CaaS	The basic CaaS principles relate to an offload of processing tasks to external compute resources. In CaaS, external compute resources can be made available to a specific entity or user device through a well-defined open interface,	D5.1
Confidential Computing	CC	Technologies ensuring that the data in use are protected against threats from malicious insiders with administrative privilege, direct access to hardware and, malwares that exploit bugs in the environment in which application runs on.	D4.1
constrained-envelope Continuous Phase Modulation	ceCPM	A phase-modulated single-carrier waveform with a controlled degree of envelope variations	D2.1
Container	N/A	A piece (package) of software within the code and all its dependencies. An evolution of a Virtual Machine.	https://www.docker.com/resources/what-container
Containerised Network Function	CNF	NF implemented by means of (one or more) containers	D6.1
Continuous learning	N/A	When new information is learnt without forgetting previous knowledge.	D4.1
Continuous Phase Modulation –	CPM-DFTS-OFDM	DFTS-OFDM where symbols are samples from a CPM modulator	D2.1

DFT spread OFDM			
Control-Communication-Codesign	CoCoCo	Joint (cross-layer) study and optimisation of control applications in I4.0 scenarios and the underlying communication service. Often extended to Control-Communication-Computation-Codesign to reflect the importance of (local) computation capabilities.	D7.1
Core Network	CN	An architectural term relating to the part of 3GPP System which is independent of the connection technology of the terminal (e.g., radio, wired)	https://www.etsi.org/deliver/etsi_tr/121900_121999/121905/16.00.00_60/tr_121905v160000p.pdf
Cramér-Rao Lower Bound	CRLB	A type of error bound (EB), based on Fisher information theory.	D3.1
Crest factor	N/A	A parameter of a waveform, such as alternating current or sound, showing the ratio of peak values to the effective value.	D2.1
CRUD Operations	CRUD	Set of functions required to manage the instantiation (creation), maintenance (update) and termination (deletion) of the network elements.	https://www.etsi.org/deliver/etsi_gs/MEC/001_099/001/02.01.01_60/gs_MEC001v020101p.pdf
Data Age of Information	Data AoI	A time-evolving measure which characterises information freshness at the receiver (e.g., at the cloud edge where an ML model is updated). The AoI at a given time instant is defined as the time difference between the focused timestamp and the time at which the observed state (or data packet) was generated.	D4.1
Data economy	N/A	Capability of achieving high inferencing accuracy with a smaller amount of learning data.	D4.1
Data poisoning	N/A	A process by which an adversary injects malicious points in the training dataset to influence the learning process and degrade the algorithm's performance.	D4.1
Data privacy protection	N/A	Data collection procedures to train the model should adhere to any regulations plus ethical obligations.	D4.1
Data quality	N/A	How useful and relevant the data are to model training - assuming the same quantity, higher quality data achieve better model convergence and flexibility.	D4.1

Decision problem	N/A	Problem requiring an algorithm to select a set of best actions provided a context.	D4.1
Deep Edge	N/A	Same as Extreme Edge	D6.1
Dependability	N/A	Dependability is the "ability to perform as and when required" [IEC61907]. Dependability consists of the attributes: availability, reliability, safety, integrity, and maintainability [ALB+04]. End-to-end dependability refers to dependability from the application perspective, encompassing multiple services (c.f. Productivity)	D7.1
DFT-spread OFDM	DFTS-OFDM	An OFDM waveform where a DFT is applied to symbols prior to their mapping onto subcarriers	D2.1
Differential Privacy	DP	Concept enabling quantifying privacy by bringing a bound on the probability that two datasets can be distinguished.	D4.1
Digital Signal Processing	DSP	The use of digital processing, such as by computers or more specialized digital signal processors, to perform a wide variety of signal processing operations	D2.1
Distributed learning with frugal AI	N/A	Distributed learning enables models to be trained without expensive communication of acquired data. Frugal AI enables learning models based on small amounts of data.	D4.1
Dynamic Function Placement	DFP	DPS the act of dynamically place network functions. This is done by deploying intelligent algorithms to orchestrate differentiated services optimally across multiple sites and clouds, based on diverse intents and policy constraints of dynamically changing environments.	D5.1
Edge AI	N/A	Concept incorporating collaborative multi-agent architectures, ML model decomposition and data parallelism principles.	D4.1
Edge Learning	N/A	New concept incorporating training of machine learning models and/or the consequent inference, on data collected and shared at the edge of the network, i.e.i.e., typically in edge cloud architectures. This paradigm requires a joint management and orchestration of communication, computation, and training/inference related parameters, to explore the best trade-off between	D4.1

		energy consumption, latency, and learning performance (e.g.e.g., model convergence, inferencing accuracy, etc.). Both standalone and federated architecture are involved	
Edge Network	N/A	Brings computation and data storage as close as possible to the location request.	D6.1
Element Management	N/A	Managing a single element, e.g., software update or configuration change for a specific device.	https://www.itu.int/rec/T-REC-M.3010-200002-I/en
End-to-end learning	N/A	Learning and optimising the transmitter and receiver jointly in a single process.	D4.1
Error bound	EB	Fundamental lower bounds on the error covariance of a parameter that is estimated (e.g., position error bound (PEB), clock error bound (CEB))	D3.1
Explainability	N/A	Ability of the AI/ML agent to provide justification for a recommendation based on model output.	D4.1
Explainable AI	XAI	The process of explaining why an AI agent performs, after an internal processing, a certain decision to the final user in understandable terms.	D4.1
Extreme Edge	N/A	Network domain where devices could be limited in computing and storage capabilities.	https://ieeexplore.ieee.org/document/8607067
Far Edge	N/A	Same as Extreme Edge	D6.1
Flexibility	N/A	Flexibility refers to the ability of the utilised technology and realised deployment to adapt to different tasks. This can be captured by the cost (monetary and resources) associated with the required change and by the complexity induced with the change (grade of re-use of components).	D7.1
Flexibility to different topologies	N/A	The ability of the network to adapt to various scenarios such as new non-public networks, autonomous networks (subnetworks), mesh networks, new spectrum, etc., without loss of performance and easy deployment. Addition of service capabilities and new services endpoints require no changes to existing end-to-end services.	D5.1
Fog Computing	N/A	A decentralised structure where networking, storage and computing resources are between the cloud and data source.	D6.1

Fog Network	N/A	Same as Extreme Edge network	D6.1
Frame Error Rate	FER	Ratio of data received with errors to total data received. Used to determine the quality of a signal connection. If the FER is too high (too many errors), the connection may be dropped.	D2.1
Full Network Automation	N/A	Full Network Automation is driven by high-level policies and rules without minimal human intervention. Networks will be capable of self-configuration, self-monitoring, self-healing, and self-optimisation	D5.1
Generalisability	N/A	AI-based models should be able to adapt to unseen scenarios and perform effectively.	D4.1
Generalised Optimal Sub-Pattern Assignment	GOSPA	An error metric between sets of detected and ground truth objects, generalising the RMSE.	D3.1
Goal-oriented communications	N/A	Paradigm in which communication is designed to guarantee the correct accomplishment of a common goal, e.g.e.g., control actions, learning and inference at the edge, etc. Therefore, performance indicators do not necessarily involve the correct reception of all collected data, but the ones (or their representation) needed to achieve a target level of effectiveness, i.e. goal achievement	D4.1
Ground truth	N/A	True state (e.g., true position of an object or UE).	D3.1
Homomorphic Encryption	HE	A form of encryption that allows the computation on ciphertext using specific operations without accessing to secret key nor requiring any decryption.	D4.1
Hybrid Network Function	HNF	NF implemented using diverse technologies (e.g., virtual machines, containers, physical elements or other).	D6.1
Inferencing accuracy	N/A	Applicable to many AI functionalities, depends on (and can be traded off for) data volume, inference latency, channel quality in data sharing.	D4.1
Integrated Access and Backhaul	IAB	The new radio (NR) technology in 3GPP that provides not only the access link to the UEUE, but also flexible wireless backhaul link connecting the base stations to the core network.	D2.1

Intent-based Networking	IBN	Technology proposing to transform a hardware-centric, manual network into a controller-led network that captures business intent and translates it into policies that can be automated and applied consistently across the network.	https://www.cisco.com/c/en/us/solutions/intent-based-networking.html
Interactivity	N/A	Interactivity is the extent to which humans contribute to changing their environment via telepresence in real-time. Interactivity is determined firstly by the performance characteristics of the computing and communication platform and secondly by adherence to agreed dependability parameters.	D7.1
Interface	N/A	Shared boundary between two functional units, defined by functional characteristics, signal characteristics, or other characteristics as appropriate (e.g., API: Application Programming Interface, UI: User Interface, webinterface: Can be API or UI, interface is bound to web protocols).	D3.1
Interpretability level	N/A	Measure of explainability, reasoning, contribution of input factors.	D4.1
Joint Transmission Coordinated Multi-Point	N/A	Coherent transmission from clusters of base stations to overcome the inter-cell interference within each cluster	D2.1
Large distributed and cooperative MIMO systems (D-MIMO)/Cell-free massive MIMO	N/A	Spread of a large number of antenna elements across the network (even in the form of single-antenna base stations), which provides enhanced coverage and reduced pathloss.	D2.1
Latency	N/A	Duration between initialisation of sensing/localisation procedure and acquiring localisation/sensing estimate. See also: update rate.	D3.1
Latency of AI/ML	N/A	AI/ ML components which support (near) real time decisions also have strict time constraints for inference or training.	D4.1
L_{bf}	N/A	The ratio, usually expressed in decibels, for a radio link between the power radiated by the transmitting antenna and the power that would be available at a conjugately matched receiver antenna input if the actual antennas were replaced by loss free isotropic antennas located in a perfectly dielectric,	D2.1

		homogeneous, isotropic and unlimited environment, the distance between the antennas being retained. If the distance **d** between the antennas is much greater than the wavelength $\boldsymbol{\lambda}$, the free-space attenuation in decibels will be: $L_{bf} = 20\log_{10}\frac{4\pi d}{\lambda}$.	
Learning struggler	N/A	Network device (end user or network infrastructure ones) having insufficient compute, memory and/or storage resources to update a local model in a timely fashion per e.g., a learning synchronisation requirement	D4.1
Lens Antenna	N/A	A microwave antenna that uses a shaped piece of microwave-transparent material to bend and focus the radio waves by refraction.	D2.1
Life-cycle Management	LCM	Set of functions required to manage the instantiation, maintenance and termination of a software component (e.g., NF or NS).	https://www.etsi.org/deliver/etsi_gr/NFV/001_099/003/01.05.01_60/gr_NFV003v010501p.pdf
Localisation (synonym: positioning)	N/A	The process of estimating the location of a connected device from sensor measurements. The location can be in 2D (horizontal plane) or 3D (including altitude).	D3.1
Location accuracy and timeliness for AI/ML	N/A	Location estimations enhanced by intelligent fusion with further models (mobility, maps, etc.) and additional data sources - time granularity to be considered jointly with location accuracy.	D4.1
Long-range wireless connectivity	N/A	Communication links at distances beyond 100 m	D2.1
Lower millimeter wave	Lower mmW	Frequency range 30 – 100 GHz	D2.1
Machine Learning	ML	The study of computer algorithms that allow computer programs to automatically improve through experience.	D4.1
Maintainability	N/A	Maintainability is the ability of a system to "be retained in, or restored to, a state in which it can perform as required under given conditions of use and maintenance" [IEC61907]. The attributes retainability and recoverability are also sometimes used.	D7.1

Mapping	N/A	Generating a map of landmarks (natural or artificial features used for navigation) based on sensing measurements.	D3.1
massive MIMO	N/A	Each base station is equipped with a large number of antenna elements and serves numerous user equipments simultaneously by means of highly directional beamforming techniques	D2.1
Microservices	N/A	Microservices are an architectural and organizational approach to software development where software is composed of small independent services that communicate over well-defined APIs. These services are owned by small, self-contained teams.	https://aws.amazon.com/microservices/?nc1=h_ls
Millimeter wave	mmW	Frequency range 30 – 300 GHz	D2.1
ML complexity gain	N/A	Implementation complexity reduction compared to a non-ML method.	D4.1
ML model complexity	N/A	Computational complexity of AI/ML models during either training or inference phases.	D4.1
ML model convergence	N/A	Related to training of the ML model. This indicates the loss function value that has been settled with increasing training epochs.	D4.1
Mobile Network Operator	MNO	Telecommunications service provider organisation that provides wireless voice and data communication for its subscribed mobile users.	https://ec.europa.eu/eurostat/cros/content/Glossary:Mobile_network_operator_(MNO)
Model Predictive Control	MPC	Control method to assist complex rule-based systems by learning a close-to-optimal control.	D4.1
Model Training Logical Function	MTLF	Function dedicated to training ML models, which can be then consumed on-demand by the AnLF.	D4.1
Monostatic sensing	N/A	Sensing, whereby the transmitting and receiving nodes are co-located.	D3.1
Multi-access Edge Computing	MEC	System which provides an IT service environment and cloud-computing capabilities at the edge of an access network which contains one or more type of access technology, and in close proximity to its users.	https://www.etsi.org/deliver/etsi_gs/MEC/001_099/001/02.01.01_60/gs_MEC001v020101p.pdf
Multiple-Input and Multiple-Output	MIMO	The use of multiple antennas at the transmitter and the receiver	D2.1

Network Domain	N/A	The highest-level group of physical entities.	https://www.etsi.org/deliver/etsi_tr/121900_121999/121905/16.00.00_60/tr_121905v160000p.pdf
Network Element	NE	A discrete telecommunications entity which can be managed over a specific interface.	https://www.etsi.org/deliver/etsi_tr/121900_121999/121905/16.00.00_60/tr_121905v160000p.pdf
Network Function	NF	A functional building block within a network infrastructure, which has well-defined external interfaces and a well-defined functional behaviour. In practical terms, a Network Function is today often a network node or physical appliance.	https://www.etsi.org/deliver/etsi_gs/NFV/001_099/003/01.01.01_60/gs_NFV003v010101p.pdf
Network Functions Virtualisation	NFV	Principle of separating network functions from the hardware they run on by using virtual hardware abstractions.	https://www.etsi.org/deliver/etsi_gs/MEC/001_099/001/02.01.01_60/gs_MEC001v020101p.pdf
Network of networks	N/A	Defined as a network that can both incorporate different (sub)network solutions as well as a network that easily (flexibly) can adapt to new topologies (same thing as Flexibility to different topologies also)	D5.1
Network Scalability	N/A	The network architecture needs to be scalable both in terms of supporting very small to very large-scale deployments, by scaling up and down network resources based on needs, e.g., varying traffic, utilising underlying shared cloud platform	D5.1
Network Service	NS	A composition of Network Functions and defined by its functional and behavioural specification.	https://www.etsi.org/deliver/etsi_gs/NFV/001_099/003/01.01.01_60/gs_NFV003v010101p.pdf
Network Service Meshes	N/A	Network service mesh is intended to support application-to-application and function-to-function communications in 6G networks and scenarios through dynamic and automated virtual network services, to be allocated on-demand, based on application requirements (similar to DFP).	D5.1
Network Service Provider	NSP	Type of provider implementing Network Services.	https://www.etsi.org/deliver/etsi_gs/NFV/001_099/003/01.03.01_60/gs_NFV003v010301p.pdf
Node Energy Efficiency	Node EE	Ratio of bitrate supported by the node when transmitting or receiving and the power consumed by the node	D2.1
Noise Figure	NF	The measures of degradation of the signal-to-noise ratio (SNR), caused by components in a signal chain..	D2.1

Non-Terrestrial Network	NTN	Satellites and other flying objects such as HAPS and UAVs.	D5.1
Optical Wireless Communication	OWC	A form of optical communication in which unguided visible, infrared (IR), or ultraviolet (UV) light is used to carry a signal. It is generally used in short-range communication.	D2.1
Orientation estimation	N/A	Estimating the 1D, 2D, or 3D orientation (e.g., roll, pitch, yaw) of a connected device.	D3.1
Out-Of-Band Emission	(OOB) Emission	Emission on a frequency or frequencies immediately outside the necessary bandwidth which results from the modulation process.	D2.1
Peak data rate	N/A	The maximum achievable data rate under ideal conditions (in bit/s), which is the received data bits assuming error-free conditions assignable to a single mobile station, when all assignable radio resources for the corresponding link direction are utilised (i.e., excluding radio resources that are used for physical layer synchronisation, reference signals or pilots, guard bands and guard times).	D2.1
Peak-to-Average Power Ratio	PAPR	The peak amplitude squared (giving the peak power) divided by the RMS value squared (giving the average power)	D2.1
Phased Array	N/A	An array of antenna elements which creates radiation patterns that can be electronically steered to point in different directions without moving the antenna	D2.1
Physical Network Function	PNF	NF implemented by means of (one or more) physical elements	D6.1
Positioning reference signal	PRS	Standardised pilot signal in time a frequency, used for ToA estimation.	D3.1
Power Added Efficiency	PAE	The overall efficiency of the power amplifier, including the effect of the gain of the amplifier and the input power. It is the ratio of the difference of output and input power to the DC power consumed.	D2.1
Power Amplifier linearity	N/A	The ability of the amplifier to produce signals that are accurate copies of the input, generally at increased power levels.	D2.1

Prediction problem	N/A	Problem that involves forecasting the likelihood of outcomes based on historical data.	D4.1
Privacy	N/A	The right of individuals to control or influence what information related to them may be collected and stored and by whom and to whom that information may be disclosed.	https://www.etsi.org/deliver/etsi_etr/200_299/232/01_60/etr_232e01p.pdf
Productivity	N/A	The fraction of time an application can operate as intended (i.e., targeted availability and reliability). Application-specific considerations can influence the achieved productivity given the availability and reliability characteristics of the underlying (consumed) services and the level of resilience of the application.	D7.1
Programmability	N/A	A framework that gives the possibility to update the program for specific features in a network entity	D5.1
Q-learning	N/A	Model-free general approach that requires no knowledge on the system to be controlled, rather just the reward (Q) function.	D4.1
Radio Network User plane latency	N/A	The contribution of the radio network to the time from when the source sends a packet to when the destination receives it (in ms). It is defined as the one-way time it takes to successfully deliver an application layer packet/message from the radio protocol layer 2/3 SDU ingress point to the radio protocol layer 2/3 SDU egress point of the radio interface in either uplink or downlink in the network for a given service in unloaded conditions, assuming the mobile station is in the active state.	D2.1
Ray path transmission loss	N/A	The transmission loss for a particular ray propagation path taking into account the antenna gains in that ray path direction. The use of this term is restricted to those cases, for example for multipath propagation, where several propagation ray paths are considered separately. The ray path transmission loss may be expressed by: $L = L_b - G_{tp} - G_{rp}$ [dB] where G_{tp} and G_{rp} are the plane-wave directive gains of the transmitting and	

		receiving antennas for the directions of propagation and polarisation considered.	
Receiver	RX	An electronic device that receives radio waves and converts the information carried by them to a usable form	D2.1
Regression problem	N/A	Problem of matching a function that outputs continuous values	D4.1
Reinforcement Learning	RL	ML technique used to learn sequences of actions that an "agent" should perform, given its state and its environment state, to maximize the expectation of reward for those sequences of action.	D4.1
Reliability	N/A	Reliability is the probability to perform as required for a given time interval, under given conditions [IEC61907, 22.104].	D7.1
Resilience	N/A	Resilience is defined as the ability of an application to react and adapt to challenging conditions by altering its behaviour to maintain dependability.	D7.1
Resilience and availability	N/A	This means that the network (architecture) shall be resilient in terms of service and infrastructure provisioning using multi connectivity, and separation of CP and UP, support of local network survivability if a subnetwork loses connectivity with another network, removing single point of failures	D5.1
Resiliency	N/A	The ability to limit disruption and return to normal or at a minimum acceptable service delivery level in the face of a fault, failure, or an event that disrupts the normal operation.	https://www.etsi.org/deliver/etsi_gs/NFV/001_099/003/01.01.01_60/gs_NFV003v010101p.pdf
Resistance to adversarial attacks	N/A	Capability to perform as intended when faced with adversarial attacks.	D4.1
Resources Orchestration	N/A	Subset of network functions that are responsible for global resource management governance.	https://www.gsma.com/futurenetworks/wp-content/uploads/2017/05/Virtualisation.pdf
Root mean squared error	RMSE	Square root of the average error norm between an estimate and the ground truth.	D3.1
Round-trip-time	RTT	Measurement of roundtrip delay between a BS and a UE. Involves a ToA estimate at each side.	D3.1

Safety	N/A	Unavailability (or degradation) of the service must not have catastrophic consequences (e.g., injuries, death) on users and environment. (Monetary) consequences resulting from safety requirements not being met can be quantified as “cost of service failure”.	D7.1
Scalability	N/A	Ability to dynamically extend/reduce resources granted to a Network Function as needed. This includes scaling up/down and scaling out/in	https://www.etsi.org/deliver/etsi_gs/NFV/001_099/003/01.01.01_60/gs_NFV003v010101p.pdf
Secure Multi Party Computation	SMPC	A computation paradigm that enables a set of parties to execute a joint computation of their sensitive data while revealing nothing except the information learned from the output.	D4.1
Security	N/A	The protection of information availability, integrity and confidentiality.	https://www.etsi.org/deliver/etsi_etr/200_299/232/01_60/etr_232e01p.pdf
Semantic communications	N/A	Communications that go beyond the common paradigm of guaranteeing the correct transmission and reception of data (irrespective of the meaning they convey), by targeting the correct interpretation of the data at the receiver	D4.1
Sensing	N/A	A sensor is any device, module, machine, or subsystem that detects events or changes in its environment. Sensing is then the operation of the sensor, in our case possibly including transmission and/or reception of signals and generation of measurements from these signals.	D3.1
Sensor fusion	N/A	Combining information (measurements or densities) from different sensors, such as radio signals, radar, lidar to obtain an improved estimate.	D3.1
Service	N/A	Distinct part of the functionality that is provided by an entity through interfaces. Note: Service and application (software/program) is often used as synonym. In our context service is used for software without UI.	D3.1
Service Based Architecture	SBA	A modular architecture introduced for 5G for the first time in which the control plane functionality and common data repositories of a 5G network are delivered by way of a set of interconnected Network Functions	D5.1

		(NFs), each with authorisation to access each other's services.	
Service Consumer	SC	An application, service, or software module that requires a service.	https://www.sciencedirect.com/book/9781558609006/java-web-services-architecture
Short-range wireless connectivity	N/A	Communication links at distance below 100 m	D2.1
Simultaneous Localisation And Mapping	SLAM	Process of jointly tracking the UE location and mapping the landmarks in the environment	D3.1
Simultaneous Wireless Information and Power Transfer	SWIPT	Technique to improve spectral efficiency, relying on the fact that the RF signal carries both energy and information, and thus energy harvesting and information decoding from the same received RF signal can be achieved.	D7.1
Sparse Neural Network	Sparse NN	Class of NN architectures exploiting the sparse activity and sparse connectivity properties.	D4.1
Spectral Efficiency	SE	The information rate that can be transmitted over a given bandwidth in a specific communication system. measured in bits/s/Hz.	D2.1
Spiking Neural Network	SNN	NNs the neurons of which imitate the behaviour of biological neurons; in an SNN, only active neurons transmit information	D4.1
Supervised learning	N/A	ML techniques used to learn the mapping from an input x to a desired output y. Those techniques require the knowledge of the expected output, making them suited for problems where data are annotated or labelled.	D4.1
Survival time	N/A	Survival time represents the time an application can continue operation without the reception of an anticipated signal/response by a consumed service.	D7.1
Synchronisation	N/A	Estimating the clock bias and drift of a connected device with respect to a reference. For multiple transmission and reception point (multi-TRP) based localisation, synchronisation means time synchronisation among TRPs.	D3.1
System Energy Efficiency	System EE	Ratio of the sum of bitrates in a system with several nodes (e.g.e.g., a base station and several users in a cell) and the sum of power consumptions of all the nodes.	D2.1

Tactile Internet	TI	A network or network of networks for remotely accessing, perceiving, manipulating, or controlling real, or virtual objects, or processes in perceived real time by humans or machines.	https://ti.committees.comsoc.org/
Tag	N/A	Unit, that enables communication, sensing and localisation (less complex device in comparison to UE often with focus on size, weight, costs and battery lifetime)	D3.1
Telepresence	N/A	Telepresence is the human experience of presence in an environment by means of a platform which exchanges data with humans via bidirectional communication links. There is a mixed reality (MR) continuum from real world to augmented reality (AR), to extended reality (XR) to virtual reality (VR).	D7.1
TeraHertz	THz	Frequency range 300 GHz – 3 THz	D2.1
Time-difference-of-arrival	TDoA	Measurement of the difference between arrival times of the first signal paths at a receiving device from two different transmitters.	D3.1
Time-of-arrival	ToA	Measurement of arrival time of a first signal path at a receiving device.	D3.1
Tracking	N/A	For localisation: continuous localisation of the same connected device over a given duration. For sensing, continuous localisation of the same target or targets over a given duration.	D3.1
Transceiver	N/A	An electronic device which is a combination of a radio transmitter and a receiver.	D2.1
Transmission loss (of a radio link) $\boldsymbol{L}$	N/A	The ratio, usually expressed in decibels, for a radio link between the power radiated by the transmitting antenna and the power that would be available at a conjugately matched receiver antenna input if actual antenna radiation patterns are substituted with no losses in the radio-frequency circuits. The transmission loss may be expressed by: $L = L_b - G_t - G_r$ [dB] where G_t and G_r are the directivity gains of the transmitting and receiving antennas, respectively, in the direction of propagation.	D2.1

Transmitter	TX	An electronic device which produces radio waves with an antenna	D2.1
Unsupervised learning	N/A	ML technique used to learn "relations" or patterns in an unlabelled input set x and provide a representation in an output space of smaller dimensionality y preserving properties on those relations.	D4.1
Update rate	N/A	Rate at which location or sensing outputs are reported. At most once per latency.	D3.1
Upper millimeter wave	Upper mmW	Frequency range 100 – 300 GHz	D2.1
Vendor	N/A	Entity that supplies SW and/or HW components.	https://www.etsi.org/deliver/etsi_gs/ZSM/001_099/007/01.01.01_60/gs_ZSM007v010101p.pdf
Vertical Industry	N/A	Companies, industries and public sector organisations operating in a specific sector.	https://www.gsma.com/spectrum/wp-content/uploads/2021/07/Mobile-Networks-Industry-Verticals.pdf
Virtual Machine	N/A	A compute resource that uses software instead of a physical computer to run programs and deploy apps. One or more virtual "guest" machines run on a physical "host" machine. Each virtual machine runs its own operating system and functions separately from the other VMs, even when they are all running on the same host.	https://www.vmware.com/topics/glossary/content/virtual-machine
Virtual Network Function	VNF	NF implemented by means (one of more) virtual machines.	D6.1
Vividness	N/A	Vividness is the characteristic richness of telepresence that depends on how the computing and communication platform represents the environment to human senses. The capabilities of the computing and communication platform are determined by the number of sensors and actuators and the sensor and actuator resolutions.	D7.1
Wireless energy transfer	WET	Radio frequency based wireless energy transfer provides energy over-the-air using so-called power beacons (PBs)	D7.1
Zero-Crossing Modulation	ZXM	A waveform that applies temporal oversampling and 1-bit quantisation at the receiver to achieve reasonable spectrum efficiency with improved energy efficiency and the relaxation of hardware requirements.	D2.1