SPONSORED BY THE

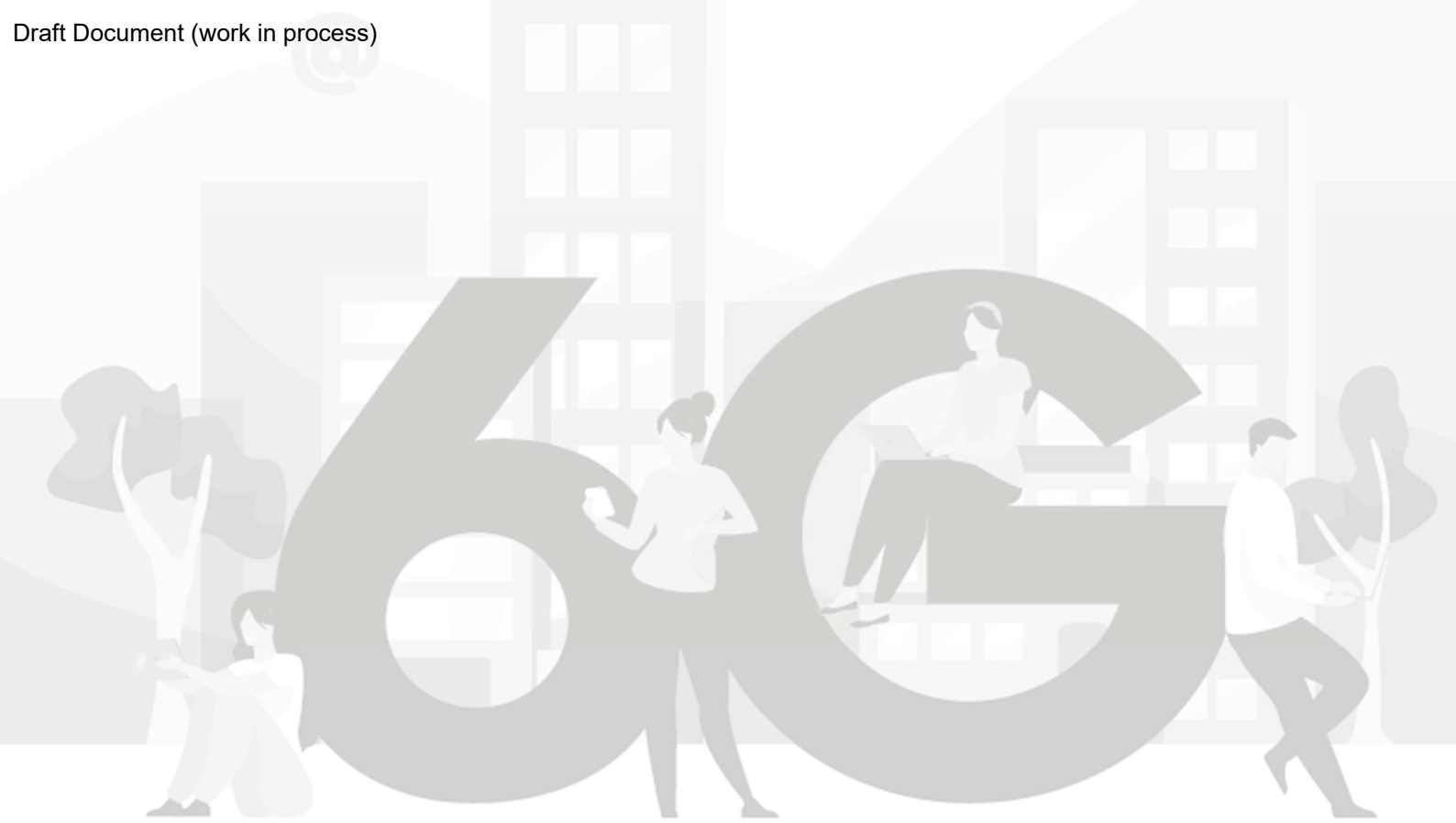
Federal Ministry
of Research, Technology
and Space

6G Platform Germany

Whitepaper

TRUSTWORTHINESS IN 6G-SYSTEMS

ESTABLISHED BY THE
WORKING GROUP TRUST



PREFACE

This White Paper was created by members of the working group “Trustworthiness” of the 6G Platform Germany, the umbrella organization of the German 6G Program, which is funded by the German Federal Ministry of Research, Technology and Space (BMBF). The objectives of the working group are to reach a common understanding related of various aspects of trustworthiness, to collect different views and needs, to understand the current R&D landscape, to identify gaps, to conduct in-depth research into 6G-oriented trustworthiness mechanisms and to articulate the German vision regarding trustworthiness for 6G. The results presented in this white paper are derived from network providers, suppliers, industries, and academic partners within the 6G Platform, which encompasses 33 German projects backed by 700 million Euro in funding. Consequently, the insights in the white paper are pertinent to industrial and academic collaborations and associations. Furthermore, political representatives and ministries involved in supporting research and industry can apply these insights to their proposed scenarios. Besides, the working group fosters the knowledge exchange and cooperation among all stakeholders interested in trustworthiness of future networks including 6G. Thereby trustworthiness is a broad concept combing many different aspects such as safety, security, privacy, resilience, reliability, assurance to name just a few. As a result, this white paper conveys the current views of the working group. It presents our own definition to trustworthy by design for 6G networks covering three directions: the devices, the services and the infrastructures. The sectors to which trustworthiness can be applied to as well as the role that standardization, certifications and regulations play in ensuring trustworthiness are also discussed. It also explores trustworthiness in 6G, defining its scope, challenges, and technical enablers. The view of the working group, as shown in this white paper gives a glimpse of a German 6G trustworthiness vision, by also managing liaisons and collaborations with other European / International 6G programs to harmonize concepts and results for joint dissemination. In consequence, we consider that this white paper serves as a valuable resource for researchers, policymakers, and industry leaders, guiding the evolution of 6G towards a future built on trustworthiness.

Editorial Board and Contributors

Editors

Organization- and Editorial Team

Authors

Contributors

Supporters

Acronyms

3GPP	3rd Generation Partnership Project
5G	Fifth Generation Mobile Networks
6G	Sixth Generation Mobile Network
AAA	Authentication, Authorisation and Accounting
ADAS	Advanced Driver Assistance Systems
ADR	Attack detection Rate
AI	Artificial Intelligence
AIRS	Aerial Intelligent Reflecting Surface
API	Application Programming Interface
AR	Augmented Reality
BSI	Bundesamt für Sicherheit in der Informationstechnik Germany (Federal Office for Information Security)
CCS-GI	Cybersecurity Certification Scheme – German Implementation
CVM	Confidential Virtual Machine
DDOS	Distributed Denial of Service
DID	Decentralized Identifier
DIN	Deutsches Institut für Normung (German Institute for Standardisation)
DLT	Distributer Ledger Technology
DoS	Denial of Service
DRL	Deep Reinforcement Learning
E2E	End-to-End
EHR	Electronic Health Record
EL	Ensemble Learning
ENISA	European Union Agency for Cybersecurity
EnWG	Energiewirtschaftsgesetz (Energy Industry Act)
ETSI	European Telecommunications Standards Institute
FL	Federated Learning
FPR	False Positive Rate
GDPR	General Data Protection Regulation
gNB	next-generation NodeB
GSMA	GSM Association
IBN	Intent-based Networking
ICAS	Integrated Communication and Sensing
IEC	International Electrotechnical Commission
IDM	Identity and Access Management
IDS	Intrusion Detection System
IoT	Internet of Things
IRTF	Internet Research Task force
ISMS	Information Security Management System
ISO	International Organization for Standardization
IT	Information Technology
ITU	International Telecommunication Union
JCAS	Joint Communication and Sensing
KRITIS	Kritische Infrastrukturen (Critical Infrastructures)
LLM	Large Language Model
LoT	Level of Trust
LoTAF	Level of Trust Assessment Function
NFV	Network Functions Virtualisation
ML	Machine Learning
MLOps	Machine Learning Operations
MNO	Mobile Network Operator
NESAS	Network Equipment Security Assurance Scheme
NGNM	Next Generation Mobile Networks Alliance
NIS2	EU Network and Information Security Directive (Second Revision)
NIST	National Institute of Standards and Technology
OECD	Organisation for Economic Co-operation and Development
OSS	Open-Source Software
PhySec	Physical Layer Security
PII	Personally Identifiable Information
PKI	Public Key Infrastructure
PQC	Post-Quantum Cryptography
QoE	Quality of Experience
QoS	Quality of Service
RAN	Radio Access Network
SAC	Soft Actor-Critic
SCF	Sensing Control Function
SEV	Secure Encrypted Virtualization
SGX	Software Guard Extensions
SLA	Service Level Agreement
SNP	Secure Nested paging
SPCTM	Sensing Policy, Consent, and Transparency Management
SPF	Sensing Processing Function
SSI	Self-Sovereign Identity
TaaS	Trustworthiness-as-a-Service
TC	Trustworthiness Characteristics
TCB	Trusted Computing Base
TDX	Trust Domain Extensions
TEE	Trusted Execution Environment
TKG	TeleKommunikationsGesetz (Telecommunications Act)
TSN	Time Sensitive Networking
UDM	Unified Data Management
UE	User Equipment
URLLC	Ultra-reliable Low Latency Communications
V2X	Vehicle-to-Everything
VC	Verifiable Credential
VDE	Verband Deutscher Elektrotechniker (Association for German Electrical Engineers)
VDI	Verein Deutscher Ingenieure (Association of German Engineers)
XAI	Explainable Artificial Intelligence
XR	Extended Reality
ZKP	Zero-Knowledge Proof
ZTA	Zero Trust Architecture

Table of Content

Introduction

Scope of the White Paper

The scope of the white paper focuses on key aspects of trustworthiness in 6G under the visions of diverse contributors such as academic researchers, telecom systems vendors and mobile network operators. The white paper examines the gap between existing trustworthiness requirements and needed developments to ensure secure, resilient, reliable and privacy-preserving 6G systems.

A clear and comprehensive definition of trustworthiness is provided, along with an exploration of trustworthy by design fundamental aspects. Key technical enablers, such as AI-driven trust management systems, confidential computing, and zero-trust mechanisms, are analysed to demonstrate how emerging innovations can enhance trustworthiness in 6G systems. The paper explores how these technologies mitigate threats and privacy concerns, ultimately fostering a more secure and trustworthy 6G ecosystem. The paper highlights also the importance of device certification for compliance with German existing and emerging certification frameworks (e.g., EU Cybersecurity Act, BSI requirements), emphasizing why trustworthiness is crucial for businesses operating in different sectors.

Additionally, the findings of the white paper are discussed in relation to standardization efforts and regulatory frameworks that govern trust in 6G, highlighting Germany's vision for the future of 6G in term of trustworthiness. The white paper is conceptual and will not delve into detailed technical implementations of trustworthiness-related measures and algorithms. It provides a high-level understanding of trustworthiness in 6G systems.

The white paper does not focus on business or commercial aspects, its goal is to provide strategic insights into trustworthiness rather than delve deeply into specific market dynamics or economic models.

Goal of the White Paper

With the upcoming 6th generation, mobile networks will become an even more important part of many critical infrastructures making 6G systems themselves an infrastructure of very high criticality. The reasons are not only that 6G systems will offer enhanced connectivity in terms of latency, bandwidth and connected devices. But 6G systems will bring new features such as Artificial Intelligence (AI) at the edge or enhanced sensing capabilities by means of integrated communication and sensing (ICAS).

Such new possibilities open the door for applications like ubiquitous robotics of all kinds. Parts of these applications were envisioned for 5G – and will become reality with 6G - but if and only if the 6G system is designed to offer the required level of trustworthiness. This is the only sustainable way to mitigate the newly arriving threats and risks associated with the evolution of mobile networks.

While in 5G data was at risk, in 6G human lives are in danger. This white paper presents an overview of attackers, attacks and threats regarding 6G systems. It highlights current pain points but presents also presents existing or envisioned solutions. In doing so it can be seen as an update of previous reports on trustworthiness in 6G reflecting new technical developments, but also reflecting the progress in standardization and regulation. We will conclude our white paper with a call for action to foster a holistic approach towards trustworthy 6G systems.

We want to point out, that although trustworthiness is a broad concept covering many different aspects of a given system. In this white paper, we concentrate on trustworthiness characteristics which have to do with security and privacy in the broadest sense. Besides the classical security goals (confidentiality, integrity, availability) the overall resilience of 6G systems is an important aspect in this regard.

This white paper was created by members of the working group trustworthiness (WG-TRUST), which was initiated by the 6G Platform Germany. The 6G Platform Germany is the German umbrella project which brings together 33 German projects backed by 700 million Euro in funding from the domain of future networks with a strong focus on 6G systems.

Structure of the White Paper

This white paper is structured as follows: in Section 2 we introduce our understanding of terms related to trustworthiness. Section 3 provides an overview of new technologies envisioned for 6G detailing the associated attacks and threats not only to these technologies but also to 6G systems. In Section 4 we introduce key enablers that can strengthen the trustworthiness of upcoming 6G systems. Although many of these enablers are technical ones we also elaborate on regulatory and organisational means including standardisation. Section 5 presents our perspective on how to leverage these enablers to mitigate threats and foster a trustworthy 6G systems. The white paper is concluded with a call for action and an envisioned roadmap related to building trustworthy 6G systems.

Chapter 2:

Definition of Trustworthiness

2

Definition of Trustworthiness

1. General

There exist many different terms related to trust and trustworthiness and many different understandings what these terms mean. To avoid ambiguities, we first define the meaning of relevant terms used in this white paper. Moreover, we relate trustworthiness to other important concepts like resilience, dependability or reliability. Afterwards we provide some insights on how trustworthiness of a given system can be measured.

2. Trustworthiness Related Terminologies

2.1 Trustworthiness

Throughout this white paper we understand trustworthiness as an objective and ideally measurable property of a given system – in contrast to trust, which is a subjective belief. This interpretation is based on definitions of the terms specified by the International Organization for Standardization (ISO) and the National Institute of Standards and Technology (NIST).

In [1], ISO defines trustworthiness as "ability to meet stakeholders' expectations in a verifiable way". NIST describes trustworthiness as follows [2] : "trustworthiness of a system is based on the concept of assurance. Assurance is the grounds for justified confidence, justified confidence is derived from objective evidence and evidence is produced by engineering verification and validation methods." In the same document NIST states: "Trust is a belief that an entity meets certain expectations". It is important to note that trustworthiness covers many different so-called trustworthiness characteristics (TCs). They can be related to security, e.g., confidentiality, integrity, availability etc. but can also be from different domains like usability, sustainability, functionality etc.

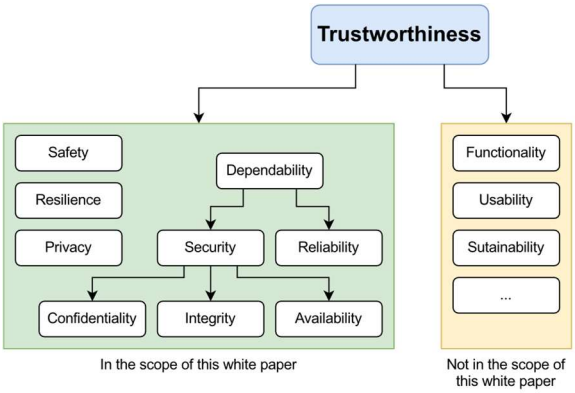


***Figure 1:** Trustworthiness covers many different so-called trustworthiness characteristics. In this white paper we concentrate on TCs which are from the general domain of IT-security and privacy.*

This understanding of the term "trustworthiness" is in line with the definition given by the European 6G flagship project Hexa-X. The project defined "trustworthiness" as the ability of the network to guarantee the confidentiality and integrity of the End-to-End (E2E) communications, while ensuring privacy and operational resilience. Note that Hexa-X only addresses security and privacy related aspects of the overall set of expectations a given stakeholder might have. The related concept "Level of Trust (LoT)" is defined by Hexa-X as the key value indicator to measure and evaluate trustworthiness of a 6G E2E environment [3].This assessment method offers the network's stakeholders insights into the security, the privacy, and the resilience of the network. In a second step and to facilitate this evaluation, the succeeding European 6G flagship project Hexa-X-II introduces the "Level of Trust Assessment Function (LoTAF)", providing a systematic framework that allows quantifying trust levels that ensures the integrity, security and reliability of the network.

The system design paradigm "trustworthiness by design" consists of embedding trustworthiness-enhancing elements during the design phase of the system to support and sustain trust among and stakeholders. In the context of 6G systems, trustworthiness by design refers to the integration of trustworthiness characteristics (TCs) to the entire lifecycle of the 6G system. This approach ensures that the 6G system keeps attributes such as security, privacy, reliability, resilience, and safety. By embedding these attributes into the 6G architecture, the 6G system will meet stakeholders' expectations, allowing confidence of users. Embedding users' trustworthiness to the design of the network can be ensured, for example, by adding security authentication mechanisms or encryption protocols during the design and development phases of 6G systems. A related concept is "native trustworthiness" which is defined as the ability of a device, infrastructure, network or service to guarantee security, privacy, reliability, safety and resilience as the key features to be built into its architecture. Trustworthiness features need to be embedded directly into the core of the network. For 6G systems, the Next Generation Mobile Networks alliance (NGMN) [4] has defined native trustworthiness as a seamless integration of trustworthiness principles into the core system, where trustworthiness-related characteristics are considered as native. This aligns with the idea that trust is "built-in" at every level and layer, allowing automatic and adaptive trustworthiness management of the system. Looking at the ongoing standardisation related to the 6G system as currently executed by the 3rd Generation Partnership Project (3GPP), a proposal for "Trustworthiness as a Service (TaaS)" was introduced [5]. The related requirements draft introduces

TaaS as a new terminology to describe the services based on trustworthiness established and managed by 6G. Potential proposed requirements include, but are not limited to, providing trust-related and security mechanisms for requests from any network consumer. Mechanisms to collect, assess and continuously evaluate trustworthiness are required to maintain and increase the confidence of 6G users. Referring to these potential requirements, trustworthiness is considered as "a measurable and consistent belief and/or confidence of the value provided by system stakeholders". Trustworthiness in 6G systems, can be measured as a "multi-dimensional and dynamic parameter, reflecting aspects of reliability, security, privacy, resilience and reputation".
These definitions and concepts are in line with our understanding of trustworthiness as introduced above and used throughout this white paper.

2.2 Trustworthiness Characteristics

Trustworthiness is a broad concept covering many so-called TCs of a given system. This white paper focuses on TCs related to the domain of security and privacy. Nevertheless, trustworthiness covers additional TCs like usability, sustainability or functionality, although such TCs are out of scope of this white paper. In the next subsections we shortly introduce definitions for TCs which are relevant for this white paper.

Security

Security refers to "the resistance to intentional, unauthorised act(s) designed to cause harm or damage to a system" [6]. This includes the preservation of confidentiality, integrity and availability of information [7].

Confidentiality

Confidentiality is "the property that information is not made available or disclosed to unauthorised individuals, entities, or processes." [8].

Integrity

Integrity is "the property that data has not been altered or destroyed in an unauthorized manner" [8]. Although this is often the primary goal of integrity, preventing unauthorized modifications is very hard if not impossible to achieve in practice. Therefore, we relax the integrity requirement stating that unauthorized modifications should be at least detectable leading to the following definition: Integrity is the property, that data has not been altered in an unauthorised manner or that unauthorised modifications can be detected.

Availability

Availability is "the property of being accessible and usable upon demand by an authorized entity" [8].

Privacy

Privacy refers to the "freedom from intrusion into the private life or affairs of an individual when that intrusion results from undue or illegal gathering and use of data about that individual" [9]. Privacy covers the "rights and obligations of individuals and organisations with respect to the collection, use, retention, disclosure and disposal of personal information." [10]. Privacy is always related to personal data/ personally identifiable information (PII). Note that privacy covers more aspects than preventing the disclosure of PII. Therefore, privacy goes beyond confidentiality.

Reliability

Reliability is defined by 3GPP in the context of network layer packet transmissions, as the percentage value of the packets successfully delivered to a given system entity within the time constraint required by the targeted service out of all the packets transmitted [11]. This is in line with the more general definition from IEC and ISO (provided in various standards), which define the reliability of a system as the "ability to perform as required, without failure, for a given time interval, under given conditions" [12]. This is very similar to the reliability definition used by ETSI [13] within their Network Functions Virtualisation (NFV) framework. Therefore, we follow the definitions by IEC/ISO/ETSI in this white paper.

Dependability

Dependability is the "ability to perform as and when required" [12]. Dependability is often seen as an umbrella term which includes other characteristics such as availability, and reliability. Note that there are other definitions of dependability which we do not use in our white paper, since these definitions contain references to trust or trustworthiness. One such example is the definition provided by IFIP Working Group 10.4: "the trustworthiness of a computing system which allows reliance to be justifiably placed on the service it delivers" [14].

Resilience

Resilience in an engineered system is commonly known as the ability to maintain an acceptable level of service while facing faults, disruptions, or unforeseen threats [15]. Adaptation and recovery in response to these challenges are distinguishing characteristics of a resilient system. A resilient network is specifically designed to withstand potential failures or significant disruptions, e.g., caused by natural disasters or (cyber) attacks, while ensuring continuous operation, even if at a reduced capacity. Key attributes of resilient networks include self-awareness and automatic reconfiguration, which may be required across different system layers. Resilience also includes the ability to learn from past experiences, thereby improving the robustness of the network against similar future challenges. ISO captures this by defining resilience as the "ability to anticipate and adapt to, resist or quickly recover from a potentially disruptive event, whether natural or man-made" [16].

Safety

Safety is the "expectation that a system does not, under defined conditions, lead to a state in which human life, health, property, or the environment is endangered" [17].

3. Relationship between Quality and Trustworthiness

There exists a strong relationship between quality, understood as "degree to which a set of inherent characteristics of an object fulfils requirements" [18], trustworthiness defined as "Worthy of being trusted to fulfil whatever critical requirements" [2] and trustworthy introduced as "the degree to which the behavior of a component is demonstrably compliant with its stated requirements" [2] exist. In the context of 6G systems, this relationship becomes more explicit. Trustworthiness can be considered as an overarching objective, encompassing security, privacy, reliability, safety as well as ethical considerations. While quality still being a foundational pillar, we adopt trustworthiness as the central term throughout this white paper, since in our view it is the more "natural" term considering the scope of this white paper. Moreover, certain quality aspects are not only relevant in the context of 6G systems but can be understood as trustworthiness characteristics. This covers e.g., quality of service (QoS) or quality of experience (QoE). Thus, quality and trustworthiness in 6G are not separate objectives but closely reinforcing dimensions that must be jointly optimized to ensure technical efficiency and societal trust in 6G.

4. Measuring Trustworthiness

Despite multiple attempts known from the scientific literature or standardisation documents, there is currently no comprehensive and sound method to measure trustworthiness. In the following we provide our view regarding concepts and aspects which should be considered when measuring trustworthiness. However, we do not claim to provide a definitive answer on how to measure it.

Amid diverse viewpoints surrounding 6G trust and trustworthiness, whether it is perceived as mere jargon symbolising an abstract concept or as a concrete engineering challenge that must be addressed in the design and deployment of 6G system, trustworthiness must be evaluated using both the objective measures of the trustworthiness characteristics as well as the expectations of the different stakeholders. The former has a close relation to Quality of Service (QoS) metrics while the later relates to Quality of Experience (QoE) metrics. While objective measures exist for assessing the trustworthiness characteristics, subjective assessments are also crucial to capture user's expectations considering perceptions and experiences. In line with this view, we defined trustworthiness as an objective and measurable property of a system, characterized by a set of attributes. However, to arrive at a quantifiable assessment of trustworthiness, it must also be evaluated by considering the trustor's expectations. In the context of 6G networks, this dual perspective means that trustworthiness is a combination of objective factors and subjective expectations. Following the trustworthiness definition ("ability to meet stakeholders' expectations"), this combination can be expressed as ratio between the measured trustworthiness characteristics and the expectations. Let us consider "***TC**" as* the set of the measured values "tc_i" for different trustworthiness characteristics and "***E**" as* the set of values "e_i" expressing the expectations of a given stakeholder, whereas e_i is the expectation with respect to "tc_i"(resp. the corresponding trustworthiness characteristic). Let "***n**"* be the size of these sets. The trustworthiness "***T***" can then be expressed as normalised sum of the ratios "r_i" between the expectations and the measured trustworthiness characteristics: $T = \frac{1}{n}\sum r_i = \frac{1}{n}\sum \frac{tc_i}{e_i}$.

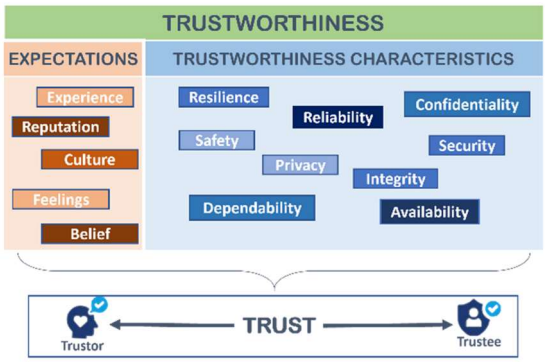
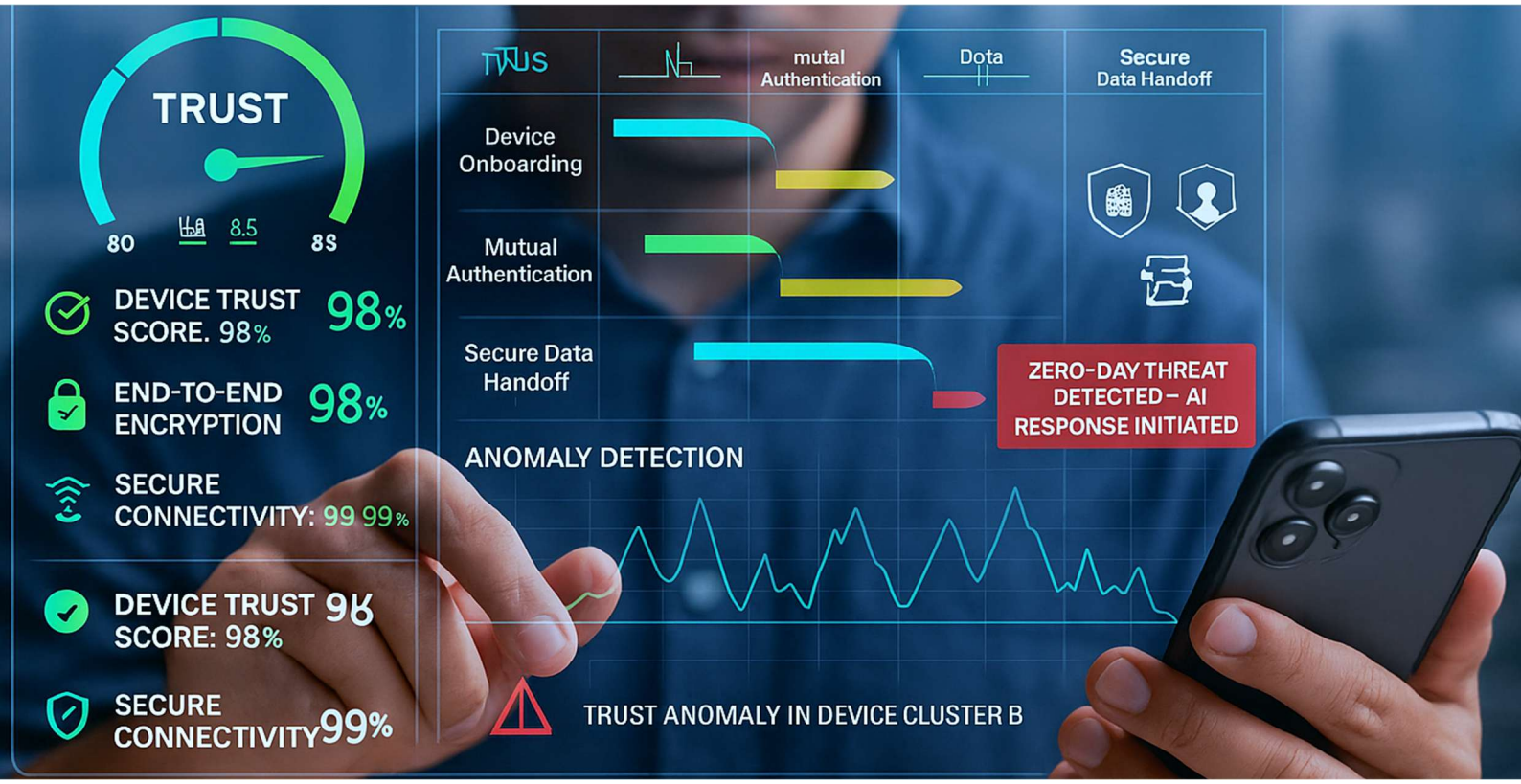


Figure 2: Trustworthiness constitutes of the objective measurable trustworthiness characteristics and the expectations of the stakeholders. As such trustworthiness is one influencing factor regarding the subjective trust decisions of humans.

Nevertheless, this simple approach leaves many questions open: How to handle the case, where a stakeholder has no expectations regarding a given trustworthiness characteristic, i.e. "$e_i=0$?" Should "$r_i \leq 1$" hold, i.e. the expectations cannot be overachieved? Would it be better to use a product instead of a sum? Should it be a weighted sum/product? Having one value for trustworthiness makes comparison easier but reduces expressiveness, would it be therefore better to have a vector, which contains the values of "r_i" as elements? While trustworthiness characteristics are measured by objective quantities, the expectations of the trustor usually incorporate subjective judgments, which are important to understand if the system's trustworthiness is perceived as sufficient. These qualitative factors include for example user experience, cultural norms, and feelings. Moreover, as different 6G use-case require different trustworthiness-related attributes and expectations, the evaluation of trustworthiness, must be flexible and context-sensitive. What is considered as "trustworthy" in a context may differ from another. A scenario-aware measurement framework is therefore essential. In order to still support some comparison of different system designs regarding trustworthiness and to allow an overall judgement, one could define some stakeholder profiles, each of which containing the expectations from the perspective of the related stakeholder. Potential profiles are: "end user", "operator", "government", and "society".



4.1 Qualitive Measurement of Trustor's Expectations

Qualitative perception can play a huge role in measuring the trustworthiness of the 6G system. It refers to understanding how users feel and perceive trustworthiness in the 6G system and that are their expectations in general. Qualitative measurement emphases on considering subjective factors that can have impact on the confidence of the users rather than numerical or statistical data. By evaluating user sentiment, transparency, beliefs, experience, cultural expectations, and reputation, stakeholders can enhance confidence in the 6G system. For example, surveys and interviews can provide insights into public perception. These factors, although not directly measurable in numerical terms, are critical in determining user confidence and satisfaction. Different factors can be used to measure and evaluate qualitatively expectations regarding 6G systems, some of them are:

- User feeling and perception: understanding how the systems' users feel about the system and its attributes is a key element. Methods such as surveys and interviews can help in revealing areas where trust may be lacking.
- Cultural Expectations: Understanding and addressing cultural differences is essential for building trust in the 6G system. Users from privacy-conscious regions may expect more stringent data protection than those in areas with more relaxed privacy norms.
- Reputation and Historical Experiences: Users usually base their trust on the reputation of the 6G system service provider or its historical performance. Generally, a positive experience from a long-term user can strengthen trust in the 6G system while a history of data breaches or service outages may reduce it. By using some tools, those qualitative factors can be measured to offer a deeper understanding of user expectation on 6G systems, some of them are illustrated in figure 3.

4.2 Quantitative Measurement of Trustworthiness Characteristics

In 6G systems, trustworthiness characteristics are mostly objective as it is grounded in measurable metrics that evaluate how well a system can uphold security, privacy, reliability, safety, and resilience under real-world conditions. 6G systems depend on well-defined and quantifiable elements like encryption, authentication, and intrusion detection systems.

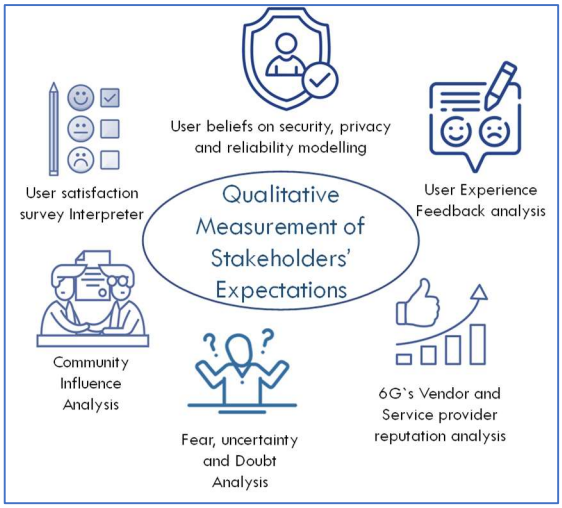


Figure 3: *Means for qualitative measurement of the expectations of stakeholders.*

These metrics help ensuring that the 6G system can withstand not only technical failures but also targeted disruptions, which are often driven by AI-enhanced attack vectors. In consequence, measuring quantitatively trustworthiness characteristics of 6G systems essentially consists of evaluating the objective metrics that define the 6G systems' such ad security, privacy, and resilience.

Privacy protection, for example, as a core element of trustworthiness in 6G, is a critical aspect in the 6G era due to the vast amount of personal data transmitted across ultra-dense networks. It can be measured and evaluated using several metrics such as Anonymization Success Rate (ASR).This measure reflects the effectiveness of anonymization techniques in preventing re-identification of data subjects. A high ASR reflects a low risk of privacy breaches and thereby enhances user trust and compliance to regulatory. Trustworthiness in 6G also hinges on resilience, the systems' capacity to anticipate, absorb, and recover from failures and malicious disruptions. This can be evaluated through some metrics like service continuity rate, disaster recovery time, fault tolerance, etc. These metrics quantify how quick and reliablly the system can recover from failures, ensuring that main users' services remain uninterrupted. Security serves as the foundational pillar of trustworthiness and many quantitative indicators are commonly utilised to assess the robustness of a 6G system's security framework as well as Attack Detection Rate (ADR), representing the percentage of malicious activities identified by intrusion detection systems and False Positive Rate (FPR) that measures the proportion of benign or legitimate actions incorrectly flagged as threats. To collect and analyse quantitative trustworthiness-related metrics in 6G, a variety of tools and frameworks are used. These tools enable systematic evaluation of key attributes such as security, privacy, and resilience. Among them:

- Security Assessment Frameworks such as Common Vulnerability Scoring System. They support mitigation strategies for risks based on objective scoring criteria.
- Privacy Risk Analysis Tools: They allow quantifying exposure levels and support privacy by design implementations.
- Network Performance Monitoring Systems allowing tracking of latency, throughput and other metrics.
- Penetration Testing and Simulation Platforms can be used to evaluate system resilience. They provide data on intrusion resistance, recovery time, and fault tolerance.
- Machine Learning-based Anomaly Detectors allow predicting trust-compromising events. Those predictive models also serve for anticipating vulnerabilities before they are exploited.

Chapter 3:

Attacks and Threats related to New Technologies for 6G systems

Attacks and Threats related to New Technologies for 6G Systems

This section will provide a non-exhaustive overview of new attacks and threads with respect to 6G systems which arise from new or extended technologies and features of 6G systems. We start with a general overview and afterwards present specific new challenges and risks.

1. General Overview regarding new risks and Threats

As telecommunications networks evolve towards 6G, the threat landscape continues to expand and become more complex. This evolution of risks can be analysed from three primary perspectives: residual risks from 5G networks, increased risks specific to 6G deployments, and the evolution of the cyber-attack ecosystem.

1.1 Residual Risks from 5G Legacy Networks

The transition to 6G inherits several fundamental vulnerabilities from 5G networks, primarily stemming from the software-defined nature of modern telecommunications infrastructure. The widespread adoption of virtualization and cloud technologies in 5G has introduced persistent vulnerabilities that continue to pose significant challenges. These risks are primarily attributed to several key factors. First, the vulnerable software components in contemporary ICT systems have become the cornerstone of software-defined, virtualized, and cloudified 5G networks. These components, despite continuous security improvements, remain susceptible to exploitation. Second, the heterogeneous nature of software resources utilized in cloud-based systems presents a significant challenge in maintaining consistent, high-level software quality across all components. This diversity in software sources and quality levels creates potential weak points in the network infrastructure. Furthermore, the increasing complexity of network operations has led to vulnerabilities in operational practices, particularly in network configuration. As networks become more sophisticated, the likelihood of configuration errors or oversights increases proportionally. Additionally, the deployment of network functions in multi-tenant clouds not dedicated to Mobile Network Operators (MNOs) introduces additional risk vectors, as these shared environments may not maintain the same level of security standards as dedicated telecommunications infrastructure.

1.2 Increased Risks in 6G Systems

The evolution towards 6G brings a significant amplification of existing risks as well as introducing new ones. This escalation is driven by several factors inherent to 6G system characteristics. The massive expansion in both the number and diversity of end-user devices, coupled with increasingly heterogeneous and complex network structures, magnifies the legacy risks inherited from previous generations. The potential divestment of responsibility across various stakeholders further complicates the security landscape, making it more challenging to maintain consistent security standards across the network. Of particular concern is the introduction of new critical use cases envisioned for 6G systems. Even when faced with threats of similar magnitude to those in previous generations, these advanced use cases present significantly higher application-layer risks due to their critical nature and potential impact on essential services. Moreover, with these envisioned critical use cases, not only the potential impact and damage increase, but thereby also the motivation of potential attackers, and hence, the likelihood of attacks rises when the stakes are higher. For example, taking down a multimedia application or some calls might have annoyed individuals, but there was no severe threat. To achieve a use case, much more targeted attacks are possible.

For instance, disturbing connected mobility could lead to accidents in the worst case; attacking a factory could stop or damage production. As risk can be measured as the product of probability and damage, both factors increase simultaneously for critical use cases. The integration of new technologies, particularly artificial intelligence, introduces additional security risks specific to 6G systems, which will be explored in detail in subsequent sections.

1.3 Evolution of the Cyber-Attack Ecosystem

The cyber-attack landscape has undergone significant evolution, presenting new and more sophisticated threats to 6G systems. This evolution is characterized by several key developments in attack methodologies and threat actors. Modern attack tools have become increasingly sophisticated, with botnets and ransomware emerging as particularly potent threats. These tools have evolved to exploit the complex nature of modern telecommunications networks, potentially causing widespread disruption across interconnected systems. A particularly concerning development is the emergence of cyber-attacks orchestrated by military organizations, governmental entities, and terrorist organizations. These actors possess significantly greater resources and capabilities compared to traditional cyber criminals, potentially leading to more devastating attacks with far-reaching consequences. Their involvement represents a substantial escalation in the potential scale and impact of cyber-attacks on telecommunications infrastructure. Moreover, the rapid pace of technological advancement in 6G systems provides new opportunities for malicious actors to abuse emerging technologies. As new capabilities are introduced to enhance network functionality, threat actors quickly adapt and develop novel attack vectors to exploit these features. The specific technologies being targeted and the methods of their exploitation will be examined in detail in subsequent sections focusing on abused technologies.

2. Threats on trustworthiness assessment

Recognizing the crucial role of trustworthiness in 6G systems, assessment mechanisms for trustworthiness (or related concepts) have become core components in end-to-end design blueprints, e.g., the Level of Trust Assessment Function (LoTAF) proposed by the Hexa-X-II project [3]. This elevation of trust assessment to a fundamental architectural element creates a significant vulnerability: the mechanisms designed to ensure trustworthiness themselves become prime targets for sophisticated attacks. This presents a paradoxical security challenge – if an adversary can compromise the trustworthiness of the trustworthiness assessment itself, the entire trustworthiness-based system becomes vulnerable. Several sophisticated attack vectors can target these assessment mechanisms in 6G systems. Data poisoning attacks can manipulate input data used for trustworthiness evaluation, causing the system to generate false trustworthiness scores. These attacks are particularly effective against machine learning-based trustworthiness evaluation systems, where carefully crafted adversarial inputs can bypass detection while corrupting outcomes. Assessment algorithm exploitation targets vulnerabilities in the mathematical or implementation aspects of trustworthiness calculation algorithms, manipulating them to favour compromised components or disparage legitimate ones. Trust anchor compromise represents another critical threat, where fundamental components inherently trusted by the system are subverted, leading to catastrophic failures in the trustworthiness evaluation chain. Man-in-the-middle attacks on trustworthiness signalling can intercept and modify communications carrying trustworthiness assessments between system components, effectively overriding legitimate evaluations. In distributed trustworthiness systems incorporating reputation mechanisms, coordinated attacks from multiple compromised entities can artificially inflate or deflate trustworthiness scores, skewing the overall assessment landscape. This “meta-trustworthiness problem” represents one of the most profound challenges in 6G security architecture: how can a system verify that its verification mechanisms are themselves reliable? The challenge is exacerbated in 6G systems due to their distributed nature, diversity of stakeholders, and varying security capabilities across different network domains. When trust decisions flow across domain boundaries, verifying assessment integrity becomes increasingly complex, particularly when domains operate under different regulatory frameworks or security standards.

3. AI requirements for trustworthy 6G

The rapid advancement of wireless communication technology has led to the emergence of 6G, the next generation of mobile networks. 6G is envisioned to revolutionize how we connect and interact with the world [19] offering unprecedented speeds, ultra-low latency, and massive connectivity. It is expected that 6G needs to offer at least 20 times more network capacity and 50 times more data transmission rate than 5G [20]. The convergence of 6G and AI will redefine the technology ecosystem, significantly boosting the performance of AI applications. However, with these advancements come new challenges, particularly in ensuring the trustworthiness of these networks. This section delves into the crucial role of AI in establishing trustworthiness in 6G communication systems. It explores the specific AI requirements, including secure AI, robust AI, and explainable AI, and examines how they enhance the security, reliability, and overall trustworthiness of 6G systems. It is envisaged that by 2030, 6G services will require a 1000× data rate and manage diverse service requirements such as massive ultra-reliable low latency communication (M-URLLC) to control autonomous entities across transport to precision manufacturing [21] [22] . Considering the evolving landscape of 6G technology, several technological trends will influence the trustworthiness of a 6G system. These include network AI, which is integral to 6G network management, optimization, and security. Open-source software and hardware will play a role in 6G development, promoting transparency and collaboration.

Virtualization and containerization will enable flexible and efficient network deployment and management.

3.1 Trustworthy requirements for AI in trustworthy 6G

AI is envisioned to be a core component of 6G systems, enabling automation, optimization, and intelligent decision-making across various network functions [20]. The International Telecommunications Union (ITU) has proposed the new 6G air interface to be AI-native and to use AI/ML to enhance the performance of radio interface functions such as symbol detection/decoding and channel estimation. Establishing an AI-native air interface for 6G systems means replacing blocks in the signal processing chain on the physical layer with trained ML models. 6G must be specified in an ontology that allows formal verification and cross-checks of implementation code, including updates [23]. This ensures that cellular infrastructures adhere to specifications. 6G systems must, therefore, have an inherent trustworthiness design to promote the effectiveness of security protection and enhance the ability of privacy protection, such as trusted computing and blockchain [24]. To ensure the trustworthiness of these AI-powered 6G systems, several key requirements need to be address.

Secure AI

Security is paramount in 6G, given the sensitive nature of the data transmitted and the critical applications it will support. Secure AI is essential to protect 6G systems and their users from various threats, including data breaches, denial-of-service attacks, and malicious manipulation of AI algorithms [25]. Secure AI in 6G includes implementing robust encryption and access control mechanisms to safeguard data from unauthorized access and tampering, protecting AI models from attacks such as poisoning, evasion, and extraction, employing AI-powered intrusion detection systems to identify and mitigate security threats in real-time [26], and adopting a zero-trust approach where no user or device is trusted by default, regardless of their location within the network.

Robust AI

Robustness refers to the ability of AI systems to maintain their functionality and performance in the face of uncertainties, errors, and unexpected events. In the context of 6G, robust AI is crucial to ensure the reliability and resilience of the network [27]. Adaptive intelligence in 6G will significantly enhance wireless connectivity by employing adaptive ML-based transceiver chains, which will power the next-generation air interface. To realize this vision, 6G is expected to learn and adapt over time, using AI-native protocols that support adaptive intelligence. Robust AI in 6G includes designing AI models that can withstand adversarial attacks and maintain their accuracy and integrity, developing AI algorithms that can effectively handle noisy data, incomplete information, and unexpected changes in the network environment, implementing mechanisms to ensure that AI systems can continue to operate effectively even in the presence of hardware or software failures, and building AI models that can adapt to dynamic network conditions and evolving user demands.

Explainable AI

Explainable AI (XAI) focuses on making AI systems more transparent and understandable to humans. In 6G, XAI will ensure transparent and secure operation at different layers of 6G systems. XAI is crucial for building trust in AI-driven decisions and ensuring accountability [28]. The need for increased explainability to enable trustworthiness is critical for 6G as it manages a wide range of mission-critical services (e.g., autonomous driving) to safety-critical tasks (e.g., remote surgery). XAI in 6G includes providing clear explanations for AI-driven decisions, enabling network operators to understand how and why AI systems are making specific choices, designing AI models that are inherently interpretable, allowing humans to understand the underlying logic and reasoning behind their predictions, enhancing trustworthiness of AI systems by providing insights into their decision-making processes and ensuring that they are aligned with ethical and regulatory guidelines, facilitating the identification and correction of errors or biases in AI models by providing explanations for their behavior, and supporting ethical and regulatory compliance by identifying and mitigating biases, promoting fairness and inclusivity in AI-driven decision-making [28].

3.2 Use Cases and Role of AI in 6G

AI is expected to transform the 6G system management and operation. Some key use cases include network management, where AI can optimize the allocation of network resources, such as bandwidth and spectrum, to meet the dynamic demands of users and applications. The terahertz spectrum and advanced modulation enable the ultra-high bandwidth and low latency needed for services like feMBB and MBRLLC [29]. AI algorithms can continuously monitor and analyse network performance, identifying and resolving bottlenecks and improving overall efficiency. AI enables 6G networks to optimise in real-time, enhancing network performance, reliability, and energy efficiency. AI can predict potential network failures and enable proactive maintenance, minimising downtime and ensuring service continuity. AI can enhance network security by detecting and mitigating threats, automating security tasks, and adapting to evolving attack patterns [29]. AI promises to enable a game-changer feature: virtualization. Indeed, this process has already started in 5G, with parts of the radio, such as base stations, going to the cloud. However, virtualizing other aspects of the network, like power amplifiers and antennas, will demand much more power and AI capabilities. With increasing virtualization, we can unlock the full power of future 6G networks, using spectrum more efficiently and cutting costs [28] (Refer to Table 1).

3.3 Risks Associated with AI in 6G

While AI offers significant benefits for 6G, it also introduces new risks that need to be carefully considered and mitigated. These include security vulnerabilities, where AI models can be vulnerable to attacks, potentially compromising the entire network's security. With faster speeds and greater digital connectedness, 6G also means greater potential for cyber-attacks and data breaches. The complexity of networks increases exponentially [30]. This complexity, while beneficial, can make systems more susceptible to sophisticated cyber-attacks, such as AI-driven threats. Potential attacks for network slicing are DoS attacks and information theft via compromised slices. Attacks on network softwarisation technologies prevent the 6G system from achieving the promised dynamicity and full automation. 6G relies on AI to enable fully autonomous networks.

Table 1.Applications of AI in 6G Network Use Cases

Use Case	AI Role	Description
Resource Allocation	Optimisation	AI can optimise the allocation of network resources, such as bandwidth and spectrum, to meet the dynamic demands of users and applications.
Network Optimization	Monitoring and Analysis	AI algorithms can continuously monitor and analyse network performance, identifying and resolving bottlenecks and improving overall efficiency.
Predictive Maintenance	Failure Prediction	AI can predict potential network failures and enable proactive maintenance, minimizing downtime and ensuring service continuity.
Security Management	Threat Detection and Mitigation	AI can enhance network security by detecting and mitigating threats, automating security tasks, and adapting to evolving attack patterns.
Virtualization	Enabling Network Flexibility	AI can enable the virtualization of network components, leading to more efficient and adaptable network infrastructure.
Sustainable Network Operation	Energy Optimization	AI can optimize energy consumption in 6G networks, contributing to environmentally friendly and sustainable operations.
Environmental Monitoring	Data Analysis	AI can analyse data from sensors deployed in 6G networks to provide insights into environmental conditions and support sustainability initiatives.

Therefore, attacks on AI systems, especially ML systems, will affect 6G. Poisoning attacks, data injections, data manipulation, logic corruption, model evasion, model inversion, and extraction are potential security threats against ML systems. Privacy concerns are heightened due to the network's ability to handle large volumes of data, raising the risk of surveillance and unauthorized data access. Attacks on collected data and the unintended use of private data lead to privacy issues as the data processing is usually not visible to the users [25]. AI models can inherit biases from the data they are trained on, leading to unfair or discriminatory outcomes. The complexity of some AI algorithms can make it challenging to understand their decision-making processes, leading to a lack of transparency and accountability. The concern is not just limited to external attacks; internal errors in AI algorithms can lead to systematic failures, disrupting network operations. Finally, the operation of AI algorithms in 6G may increase power consumption [30].

Mitigating the Risks

Several mitigation strategies can be employed to address the risks associated with AI in 6G. These include implementing strong encryption, access control, and intrusion detection systems to protect AI models and the network from attacks [26]. In the ever-evolving landscape of cybersecurity threats, employing state-of-the-art encryption methods is paramount to safeguarding data transmission within the network. 6G networks must adopt advanced encryption techniques such as quantum-resistant cryptography to ensure data confidentiality and integrity in the face of evolving cyber threats and employing techniques such as differential privacy and federated learning to protect user data while still enabling AI-driven functionalities and developing methods to detect and mitigate biases in AI models, ensuring fairness and equity in their applications [30] and promoting the use of XAI techniques to make AI systems more transparent and understandable, increasing trustworthiness and accountability. While each security measure encryption, access control, and intrusion detection systems - offers valuable protection, they also have limitations. Encryption can be computationally expensive, access control can be complex to manage in a dynamic network environment, and intrusion detection systems may not always be able to detect novel attack patterns. Therefore, a multi-layered security strategy that combines these different approaches is essential to provide comprehensive protection for AI in 6G [23].

Ethical Considerations

The use of AI in 6G raises important ethical considerations that must be addressed. These include ensuring that the collection and use of user data by AI systems comply with privacy regulations and ethical guidelines. One of the primary ethical issues concerning AI and business communications is privacy and personal data. Guaranteeing that AI algorithms are free from biases and do not discriminate against any group of users. Providing clear explanations for AI-driven decisions and ensuring that there are mechanisms for accountability in case of errors or unintended consequences [26]. Considering AI's broader social and economic impacts in 6G and ensuring that it is used to promote human well-being and societal benefit. Ensuring that AI in 6G is used for ethical purposes and does not contribute to harmful activities or social problems. Addressing the ethical implications of AI in cybersecurity, such as the potential for AI-powered surveillance and the use of AI in offensive cyber operations. The increasing use of AI for environmental awareness in 6G, while offering valuable benefits for sustainability, also raises ethical concerns about data privacy and potential surveillance. The extensive sensor deployments and data collection capabilities of 6G networks, combined with AI-powered analysis, could potentially be used for surveillance purposes, raising questions about the balance between environmental monitoring and individual privacy rights [27].

4. Risks and Threats related to Integrated Sensing and Communications

Integrated Communication and Sensing and (ICAS) systems in 6G systems integrate sensing and communication within a unified framework. Thereby the sensing part relates mainly to radar sensing. While this integration offers significant advantages, it also induces various security and especially privacy risks. The related threats can be categorised based on their impact on different layers of the system, ranging from physical-layer attacks such as jamming to sophisticated AI-driven adversarial attacks. Addressing these security and privacy challenges requires a comprehensive approach that includes robust mitigation strategies and advanced security frameworks.

In the following sections we first give a general overview of the risks and threat landscape and subsequently provide more details for some of the mentioned attacks.

4.1 Security Risks Landscape in Integrated Sensing and Communications

One of the most challenging aspects are the privacy risks associated with the new radar-based sensing features. While radar is often considered to be less privacy invasive e.g. compared to cameras it still can have severe impact on the privacy of the sensed humans. Depending on the final 6G system design, the related radar capabilities in terms of resolution and range, as well as the specific sensing use-case there exists a wide range on potential threats to privacy. Some examples are the risk of reidentification, i.e. tracking of persons, the identification of persons i.e. by means of radar-based face recognition, or the derivation of medical information like the heart rate. Since some of the information which can be derived from the (raw) radar measurements is personal data, the collection and processing of the radar data needs to be compliant to privacy and data protection regulations, e.g. the European General Data Protection Regulation (GDPR). The privacy aspects become especially challenging, since essentially every human being can be affected — especially bystanders, who are not users of the 6G system. Moreover, some of the usual approaches for achieving privacy compliance like informed consent by the data subject are not feasible.

Another pressing threat from the security domain in ICAS systems is related to Denial of Service (DoS) attacks, which can occur at both the physical and network layers. Physical-layer DoS attacks, such as jamming, are highly feasible and can severely disrupt sensing and communication signals, leading to degraded performance and potential system shutdowns. Countermeasures such as spread spectrum techniques, adaptive beamforming, and Aerial Intelligent Reflecting Surface (AIRS) solutions have been proposed to mitigate these attacks [31] [32] [33]. Similarly, network-layer DoS attacks overload network resources, causing data transmission delays and reduced situational awareness. Intrusion detection systems, rate limiting, and AI-based anomaly detection mechanisms have been suggested as effective countermeasures [34] [35].

In addition to DoS attacks, ICAS systems are vulnerable to more complex attacks, such as the Jellyfish attack, which delays or provides incomplete sensing data, impacting time-sensitive applications like autonomous driving and drone communications. Trust-based reputation systems and machine learning-based anomaly detection have been recommended as countermeasures [36]. Another concerning attack is the Intelligent Cheater attack, where malicious nodes manipulate or falsify sensing and communication data to gain unfair advantages in resource allocation. Blockchain-based verification and cryptographic authentication methods can mitigate such threats [37]. Sybil attacks and data injection attacks pose additional security risks in ICAS environments. Sybil attacks involve the creation of multiple fake identities to disrupt consensus-based decision-making, leading to the propagation of false sensing data [38]. Public key infrastructure (PKI)-based authentication and blockchain-based identity verification are promising solutions to counteract these attacks [37]. Data injection attacks, on the other hand, involve fabricating traffic or safety messages to mislead decision-making, resulting in traffic disruptions or hazardous scenarios. AI-based anomaly detection and secure data provenance tracking can help mitigate these risks [39] [40].

Other significant threats include sensor spoofing, camouflage attacks, and sensor obstruction. Sensor spoofing can mislead object detection mechanisms by introducing fake radar reflections, while camouflage

attacks can hide real objects from detection using stealth techniques. Multi-modal sensing, secure radar waveform design, and adversarial-resistant AI models have been proposed as countermeasures. Additionally, physical obstruction or blinding attacks, where attackers block or overwhelm sensors with high-power signals, can severely impact environmental awareness. Redundant sensing and adaptive power control strategies are essential in mitigating such threats.

Finally, AI-driven transformation attacks, such as adversarial manipulations of scaling, rotation, or translation, can lead to misinterpretations of object locations and orientations. These attacks highlight the need for robust AI training methodologies and adversarial machine learning defences.

4.2 Privacy Vulnerabilities and Risk Mitigation in ICAS Systems

ICAS technologies are pivotal in advancing autonomous vehicles and drones, enhancing navigation, safety, and situational awareness. However, the integration of communication and sensing functionalities introduces significant privacy threats that necessitate comprehensive mitigation strategies.

Eavesdropping on Communication Data by Radar Targets

In ICAS systems, a common approach involves designing waveforms that enable the transmitter to simultaneously communicate with downlink cellular users while tracking radar targets via reflected echoes. However, radar targets may also act as eavesdroppers, intercepting information intended for legitimate users. To mitigate these data privacy risks, physical layer security techniques can be employed. These methods prevent radar targets from exploiting communication data embedded in the ICAS waveform. Strategies include reducing the signal-to-interference-plus-noise ratio (SINR) at the target's location through artificial noise, leveraging constructive interference, and utilizing beamforming to enhance security [31] [41].

Unauthorized Localization of Radar Transmitters and Targets by Communication Users

In ICAS systems, while concerns often focus on radar targets intercepting communication data, it is equally important to consider the reverse scenario: communication users potentially estimating the location of radar transmitters. Given that ICAS waveforms serve dual purposes, facilitating both communication and sensing, communication receivers equipped with advanced signal processing capabilities can analyse received signals to infer the position of radar transmitters. By examining parameters such as time of arrival (TOA) and angle of arrival (AOA), these receivers can estimate the transmitter's location. This capability, while beneficial for network coordination and interference management, raises security and privacy concerns, especially if unauthorized users localize sensitive radar installations. To mitigate such risks, implementing physical layer security techniques, including artificial noise generation and beamforming strategies, can help obfuscate the transmitter's exact location, thereby protecting the system from potential exploitation [28] [42].

Unauthorized Collection of Location Data

Autonomous vehicles and drones continuously share location data for navigation and collision avoidance. Unauthorized parties can intercept this data to track vehicle movements, leading to privacy violations. In fleet management systems, companies may collect excessive location data beyond operational needs, raising privacy concerns. Implementing robust encryption protocols and access controls is essential to protect location data from unauthorized access. Additionally, adopting data minimization principles can ensure that only necessary location information is collected and retained [43].

Excessive or Granular Tracking

Continuous high-precision tracking of vehicles and drones can lead to user behaviour profiling, surveillance risks for private or corporate fleets, and user privacy loss. To mitigate these risks, it is crucial to establish clear data retention policies and provide users with control over their data. Implementing anonymization techniques can also reduce the risk of sensitive information exposure [43].

Unauthorized Secondary Use and Data Sharing

Vehicle manufacturers or communication providers may share collected location and sensing data with third parties without user consent, leading to privacy violations and potential misuse. Establishing transparent data-sharing policies and obtaining explicit user consent is vital to address these concerns. Additionally, implementing data anonymization and aggregation techniques can reduce the risk of sensitive information exposure [43] [44].

Inference of Sensitive Information

Even if exact location data is not leaked, attackers can analyse movement patterns to infer user habits, drone delivery patterns, and fleet operation schedules, leading to privacy breaches. Employing advanced data analytics and machine learning algorithms can help detect and prevent such inferences. Additionally, implementing robust access controls and continuous monitoring can help detect and mitigate unauthorized data access.

Lack of User Control and Transparency

Users often lack visibility into who collects their data, how it is stored or shared, and how long it is retained, eroding trust in ICAS-based mobility systems. Enhancing transparency through clear privacy policies and providing users with control over their data are essential steps in building trust. Implementing user-friendly interfaces for data management can empower users to make informed decisions about their data [43] [44].

Drone Surveillance and Data Misuse

Drones equipped with ICAS capabilities can be used for unauthorized surveillance, affecting individual privacy, corporate security, and government operations. Implementing geofencing and no-fly zone technologies can prevent drones from entering sensitive areas. Additionally, establishing strict regulations and oversight mechanisms can deter unauthorized surveillance activities.

Legal and Ethical Challenges in ICAS Mobility Systems

The lack of clear legal frameworks for ICAS in autonomous vehicles and UAVs leads to ambiguities in data ownership, ethical concerns about constant tracking, and challenges in cross-border data sharing. Developing comprehensive legal frameworks that address data ownership, user consent, and cross-border data sharing is crucial. Engaging stakeholders in the development of these frameworks can ensure that they are comprehensive and address all relevant concerns [43].

Conclusion

Addressing these privacy threats requires a multi-faceted approach, including robust security measures, stringent privacy protections, and effective management practices. Implementing privacy-preserving technologies, such as data anonymization and secure data storage solutions, can significantly enhance user privacy in ICAS systems. Additionally, fostering collaboration among industry stakeholders, policymakers, and users is essential to develop and enforce standards that protect privacy in the evolving landscape of autonomous vehicles and drones.

5. Intent-Based Networking

Intent-Based Networking (IBN) promises to greatly aid in the realization of autonomous networking, easing management and administration of services and infrastructure by providing a more user-friendly interface [45]. This relies on the concept of operator Intent which expresses the high-level goals for the network operation. An administrator merely needs to express such an intent, and the IBN should autonomously alter its configuration to meet the requirements as good as possible. To interpret and implement the operator's intent, IBNs can employ AI and ML techniques. For example, the IBN would use AI/ML for tasks already mentioned in 3.3.3 such as Resource Allocation, Network Optimization and Predictive Maintenance. Consequently, the security, and therefore the trustworthiness, of IBN is greatly affected by its reliance on SDN, AI and ML [46].

In IBN, trustworthiness of the components that interpret and implement the intent is a critical basis for establishing the trustworthiness of the system. Can the network fulfil the intent? Are intents "understood" correctly? Does the system make the correct decisions and trigger the correct actions based on its understanding of a given intent? How can this be verified? Since intents are a high-level concept, they may be open to differing interpretation. If the user's expectations of the IBN are not met, is the system still trustworthy? Furthermore, the intent handling components of the IBN need the ability to verify trustworthiness at different layers. They need to verify trustworthiness amongst each other and the underlying infrastructure, from management and orchestration down to individual hardware. This can only be provided by transitive trustworthiness.

While intents can potentially be used to express security, reliability, quality and other operator requirements, IBNs are faced with the threat of potentially malicious intents causing disruptions to the system: "Due to the inherent flexibility in expressing service or resource requests in natural language, users' statements may intentionally contain ambiguities, leading to potential network sabotage, denial of services, data breaches, and privilege escalations" [47].

Thus, the intent's and the "intenting" entity's trustworthiness needs to be ensured as well. User authentication and authorization should be a given, but the intent engine also needs to verify that the intents it receives are not malicious, in case of a compromised user or account.

6. Threats and challenges related to Quantum technologies

6.1 Quantum Computing and its Implications

Quantum computing, with its immense computational capabilities, has the potential to disrupt current cryptographic standards that are foundational to trustworthiness in 6G systems. Unlike classical computers, which perform calculations sequentially, quantum computers leverage superposition and entanglement to process information exponentially faster. This capability is poised to revolutionize fields like optimization, AI, and materials science.

However, the same power makes quantum computing a formidable weapon in the hands of adversaries.

The most significant threat arises from quantum computers' ability to break widely used asymmetric cryptographic algorithms such as RSA, Elliptic Curve Cryptography (ECC), and Diffie–Hellman (DH) key agreement protocols, which are critical for securing digital communication. Algorithms like Shor's algorithm can efficiently solve the integer factorization and discrete logarithm problems underpinning these encryption schemes. This means that any encrypted data intercepted today could be decrypted in the future once quantum computers reach a sufficient level of maturity—a threat often referred to as the "harvest now, decrypt later" strategy. This capability undermines the confidentiality and integrity of sensitive data, eroding trust in communication systems.

To ensure trustworthiness in the 6G era, it is imperative to transition to cryptographic methods resistant to quantum attacks. However, this transition must be carefully managed to avoid vulnerabilities during the migration phase, such as insecure hybrid systems or misconfigurations.

6.2 Quantum Secure Communication and Its Risks

Quantum secure communication, particularly through quantum key distribution (QKD), has been hailed as a revolutionary approach to achieving virtually unbreakable encryption. By leveraging the principles of quantum mechanics, such as the no-cloning theorem, QKD enables secure key exchange, with any eavesdropping attempt detectable in real-time. While this technology holds promise for enhancing trust in 6G systems, it is not immune to misuse. Adversaries could exploit QKD networks to create clandestine, tamper-proof communication channels, making it harder for law enforcement or cybersecurity professionals to monitor malicious activities. Additionally, vulnerabilities in QKD implementations—such as side-channel attacks—could be exploited by sophisticated attackers to compromise key exchanges, bypassing the supposed invulnerability of quantum communication systems. The trustworthiness of QKD systems also depends heavily on the integrity of their hardware. Supply chain attacks, where malicious actors introduce backdoors or vulnerabilities during manufacturing, could undermine the security of quantum communication networks before they are even deployed.

Post-Quantum Cryptography: A Critical Safeguard with Challenges

Post-quantum cryptography (PQC) represents a cornerstone of the defence strategy against quantum-enabled attacks. These cryptographic algorithms are designed to resist the computational capabilities of quantum computers, ensuring that data remains secure even in a post-quantum world. However, while PQC strengthens trust in the long term, its deployment introduces transitional risks that could be exploited by adversaries. The adoption of PQC requires careful consideration, as hasty or poorly planned implementation may inadvertently weaken security. Hybrid cryptographic systems, which combine classical and quantum-resistant algorithms, are often deployed during the transition to PQC. However, these systems can introduce vulnerabilities if not designed and configured correctly. Additionally, attackers may exploit the complexity of PQC algorithms to mount new types of attacks, such as exploiting bugs in implementations or overwhelming systems with computationally expensive operations. To maintain trust, rigorous standardization processes and careful deployment strategies are essential.

Broader Risks to 6G from Quantum Technologies

Beyond their specific implications for cryptography, quantum technologies present broader risks to the trustworthiness of 6G networks. For instance, adversaries could use quantum computers to enhance adversarial AI models, automating the discovery of vulnerabilities in 6G systems. Quantum-enhanced algorithms could also improve the efficiency and success rate of cyberattacks, such as breaking authentication protocols or evading detection mechanisms. Furthermore, as 6G increasingly integrates quantum sensors and quantum-enhanced communication systems, these components become new attack surfaces. Adversaries could target quantum-sensitive systems to disrupt services or manipulate data, undermining the reliability and accuracy of 6G-enabled applications, such as autonomous vehicles or smart cities.

Ensuring Trustworthiness in the Face of Quantum Threats

To safeguard trustworthiness in 6G networks, the risks associated with quantum technologies must be addressed through a multifaceted approach. This includes accelerating the adoption of quantum-resistant cryptographic solutions, ensuring the secure implementation of QKD systems, and maintaining rigorous oversight of quantum hardware and software supply chains. Additionally, the development of threat monitoring and mitigation strategies tailored to quantum-enabled attacks will be critical.

7. Non-Terrestrial Networks

Non-terrestrial networks (NTNs, especially satellite-based networks) are currently gaining attention and most likely will become an integral part of 6G systems. They can bring certain benefits like increased coverage as well as increased resilience e.g. in case of (large-scale) natural disasters or (physical) attacks on the ground. Nevertheless, the integration NTNs also induces new security and privacy risks and threats. With respect to security threats arise from the fact, that satellite signals cover a potentially large area, even if technologies like beam forming are applied. This could allow attackers to eavesdrop on connections of many UEs with a comparable lower number of antennas (compared to the terrestrial cellular system). Jamming is another threat to NTN. Depending on the assumptions (e.g. available resources and skill) and goals with respect to the attacker jamming of NTN connections can be seen as more difficult or easier compared to the jamming of terrestrial base stations. Given the goal of enhanced resilience with the help of NTNs one should consider powerful attackers, which have the capability to either operate strong terrestrial jammers or even able to place jammers in space. Similarly, the (physical) resilience of the satellite providing the NTN can be judged from two perspectives. One the one hand one could argue that satellites are better protected from physical attacks compared e.g. to terrestrial base stations simple because it is much more difficult to reach them for executing physical attacks. One the other hand defending a satellite with respect to physical attacks or repair/replacing it after some successful attack involves also much more effort and usually takes a considerable amount of time. This needs to be especially considered, if NTNs provide the only connectivity for certain UEs. This could e.g. involve use cases from the IoT domain there e.g. (offshore) windmills can only be controlled with the help of satellite connections. Besides security NTNs could also pose new threats to privacy. Thanks to the larger area covered by a single satellite in combination with beam forming and beam tracking could allow the attacker to track the locations of UEs more easily (compared to terrestrial networks).

Chapter 4:

Key Enablers for Trustworthiness in 6G

4

Key Enablers for Trustworthiness in 6G

In the following section we highlight selected enablers which provide helpful build blocks to realise a 6G-system which is “trustworthy by design”.

1. Decentralized Identity and Access Management

With the rise of network virtualization, service-based interactions and automation, not only humans become subject to identity management, but also technical and virtual components. Additionally, centralized identity management systems combined with the ever-increasing number of different platforms and services result in users being associated with a high number of complex identities, and central providers posing a valuable target for data breaches and identity theft [48].

1.1 Self-Sovereign Identity

Self-Sovereign Identity (SSI) poses a form of digital identity management, which differs from traditional centralized/service-based approaches by enabling the subject of the identity to manage it on their own behalf [49]. SSI, prominently based on Distributed Ledger Technology (DLT), could facilitate the delegation of specific subscriber management tasks like access management from centralised algorithms to end-user devices in the context of 6G. This shift alters the roles of various network components, allowing for more adaptable network architectures while strengthening user privacy. Infrastructure resource-sharing models can be utilised to improve network coverage in remote regions [50], taking advantage of flexible billing systems and decentralised authentication methods. A unified decentralised identity management framework could serve as a link between multiple authentication providers and applications, ensuring that only the required personal data is disclosed. As the threat of malicious deepfake usage grows, universally accessible authentication solutions are becoming increasingly crucial for verifying communication endpoints and thereby ensuring trustworthiness throughout the network.

1.2 Self-Sovereign Identity for Trustworthiness in 6G Networks

The transition from monolithic mobile networks to the highly interconnected, multi-stakeholder ecosystems of 6G introduces unprecedented opportunities and challenges, particularly in establishing trustworthiness among diverse actors such as mobile network operators (MNOs), virtual network providers, IoT device operators, and emerging applications like connected vehicles and drones. Trustworthiness in 6G—encompassing security, privacy, reliability, and scalability—requires a fundamental shift in identity and access management (IDM). SSI emerges as a key technical enabler to build this trust foundation, offering a decentralised, user-centric approach that contrasts with the limitations of traditional centralised Public Key Infrastructures (PKIs). By empowering entities—whether human users, devices, or software agents—to control their digital identities, SSI fosters secure, transparent, and efficient interactions across the heterogeneous 6G landscape.

SSI's Role in Trustworthiness

- Decentralised Trust: SSI uses Distributed Ledger Technology (DLT) to eliminate single points of failure, enabling secure, intermediary-free authentication (e.g., drones verifying with infrastructure).
- Privacy: Zero-Knowledge Proofs (ZKPs) allow selective data sharing, aligning with GDPR and boosting user confidence.
- Scalability: Verifiable Credentials (VCs) and Decentralised Identifiers (DIDs) support dynamic authentication in growing ecosystems like IoT or cross-border logistics.

2. Anomaly Detection

6G is expected to further deepen the connection between the real and digital world and increasingly support safety-critical applications. This, together with trends as network cloudification and multi-vendor systems enabled by Open Radio Access Network (Open RAN), makes monitoring the network and its components more important than ever [51] [52]. Continuous and thorough monitoring allows for the detection, identification, and response to anomalies that might threaten agreed service level agreements (SLAs) or disrupt the overall network operation. However, with the evolution of mobile networks, it is also expected that the means for network monitoring will develop. Already existing tools will be complemented by envisioned enablers like digital twins (of the network) and artificial intelligence, which are anticipated as an integral part of 6G and can help in the automated supervision of the network. Accordingly, they are expected to help detect, identify, and treat anomalies in the network. Those anomalies are not necessarily intended, i.e., induced by an attacker. They can also occur accidentally by misconfiguration, faults, or any external disturbance. However, in both cases, an appropriate reaction is required to ensure the provision of critical services. Furthermore, different parts of the network from the radio interface up to the core network can be affected and should, therefore, be observed. Current research mainly addresses the detection aspects. Recent works include the detection of anomalies and intrusion directly in the spectrum [53] [54] or on packet-level [51] [55] [56]. Whereas the former refers to jamming attacks or (unexpected) interference, the latter one considered different types of (distributed) denial of service (DDoS/DoS) attacks in private campus networks [55] and O-RAN [51]. As shown in these works, both artificial intelligence (AI) and digital twins (DT) can play different roles in the detection tasks. For instance, the radio spectrum in [53] [54] is modelled as a digital twin channel (DTC) and then compared against the actual measurements with classical approaches. Herein, AI can help in reducing the complexity of the DTC. On the other hand, the authors in [51] use DTs as a tool to train and test AI algorithms, which are then used for the detection. The study in [55] did not utilise DTs at all, but relied on the AI approaches for anomaly and intrusion detection. Even though the authors have shown good detection performance, they concede that such training-based methods with a limited number of training patterns may be prone to zero-day attacks, leaving room for future research. In addition to the spectrum and the RAN, the core can also be affected by attacks. The detection of specific anomalies has been investigated, for example, in [57]. Regarding suitable reactions, digital twins and artificial intelligence can help in the decision-making. For instance, if not too time-critical, several countermeasures could be simulated within the digital twin before being applied to the real-world network. In addition, digital twins could be used to train such AI algorithms or to test different attack scenarios, similar to the aforementioned detection method. However, although the detection of anomalies in mobile networks has received significant attention, how these anomalies are handled is rarely discussed and offers an opportunity for substantial research contributions. However, even though AI has a huge potential for anomaly detection, as shown in the existing and future work, the specific risks AI itself may introduce (cf. sec 3 in Chapter 3) should always be taken into account.

3. Zero Trust mechanisms in 6G

As 6G networks promise unprecedented levels of connectivity, performance, and autonomy, they simultaneously expose a vastly expanded attack surface due to the proliferation of intelligent devices, decentralised infrastructures, and dynamic network topologies. Traditional perimeter-based security models are insufficient in this new landscape. Zero Trust Architectures (ZTA) emerge as a critical paradigm shift to secure 6G by enforcing the principle: "never trust, always verify." The foundational framework for Zero Trust is outlined in the NIST Special Publication 800-207 [58], which defines ZTA as a security model that eliminates implicit trust in any element, node, or service, regardless of its location within or outside the security perimeter. However, it should be noted that "Zero Trust" does not mean, that one does not need to trust any component any longer. Instead, the concept of zero trust allows to reduce the number of trusted components and makes the trust assumptions more explicit. Yet there are still many assumptions on which the overall trustworthiness of the system is based. While this model has seen significant adoption in enterprise IT, its application to 6G systems remains in the early stages of research. The current body of work tends to focus on adjacent technologies—such as AI-based anomaly detection, blockchain authentication, and fine-grained access control—without yet consolidating into a unified framework for Zero Trust in 6G. As such, this section synthesises foundational principles, enabling technologies, and the implications of ZTA in the 6G context, while acknowledging the evolving nature of this field.

3.1 Principles of Zero Trust in 6G

To effectively apply Zero Trust in 6G, it's critical to understand its foundational principles.

These core ideas—continuous verification, least privilege, and assumption of breach—form the conceptual bedrock for designing secure, resilient, and dynamic security postures. In the context of 6G, these principles must be adapted to address the decentralised, heterogeneous, and high-mobility characteristics of next-generation networks.

Continuous Verification

In 6G environments, where endpoints are constantly moving and dynamically reconfiguring, continuous verification of identity, device posture, and contextual attributes is essential. Unlike static authentication mechanisms, ZTA requires ongoing assessments throughout the session lifecycle to mitigate lateral movement and session hijacking risks [59].

Least Privilege

Implementing least privilege means giving entities—users, devices, or services—only the access rights they need at the moment they need it. In 6G, this concept becomes more granular due to the microservice and network slicing architectures, where fine-grained access control is crucial [60].

Assumption of Breach

ZTA assumes that breaches are inevitable or have already occurred. This encourages proactive threat detection and response strategies, integrating security deeply into the communication fabric of 6G [61]. The assumption of breach principle necessitates segmentation, isolation, and real-time forensic capabilities.

3.2 Key Technologies for Implementing Zero Trust

Zero Trust is not a single technology but a framework that must be operationalised through a constellation of supporting technologies. In 6G, this includes leveraging AI for anomaly detection, DLT and SSI for decentralised identity, and intelligent orchestration tools for real-time policy enforcement. This subsection explores the enabling technologies that make Zero Trust both feasible and effective in the highly dynamic and distributed environment of 6G.

AI-driven Anomaly Detection

Machine learning models, especially deep learning and federated learning approaches, can detect deviations from normal behaviour in real time. In 6G, where traffic volumes and heterogeneity are high, AI becomes indispensable for scalable anomaly detection. Nevertheless, it needs to be carefully ensured, that the AI mechanisms do not induce a new attack surface [61].

DLT- and SSI-based Authentication

Decentralised authentication using DLT and SSI enhances identity integrity and auditability. In 6G, DLT and SSI can facilitate trustless identity attestation across distributed edge and fog nodes, reducing reliance on central authorities [62].

Intelligent Access Control

Context-aware access control policies, powered by AI and multi-attribute decision-making, enforce real-time zero trust decisions. For example, access to network slices or specific services can be restricted based on location, device state, and historical behaviour.

Security Orchestration

Automated security orchestration enables dynamic enforcement of ZTA policies across highly elastic 6G networks. It involves integrating SDN/NFV with threat detection systems for real-time response to policy violations [63].

Cyber-threat Intelligence

Zero trust in 6G benefits from global and local threat intelligence feeds that inform adaptive access decisions.

Threat intelligence systems contribute to shared situational awareness and allow pre-emptive defence mechanisms [64].

3.3 Benefits of Zero Trust in 6G

Adopting Zero Trust in 6G promises to deliver a range of strategic benefits. These go beyond security to encompass resilience, operational flexibility, and long-term risk reduction. This subsection highlights how these advantages align with the goals and performance expectations of 6G networks, especially in mission-critical and data-sensitive applications.

Enhanced Security

By eliminating implicit trust and enforcing rigorous verification, ZTA significantly reduces vulnerabilities associated with credential misuse and insider threats in 6G networks [65].

Improved Resilience

Zero trust supports resilient communication by containing breaches and minimizing lateral threat propagation. This is especially vital for critical 6G applications like telemedicine and autonomous transport systems [66].

Increased Flexibility

ZTA complements the dynamic nature of 6G by enabling secure, on-demand access across multi-domain environments, including satellites, UAVs, and edge nodes [67]. Reduced Risk: By minimising the trust radius, ZTA reduces the impact radius of potential compromises, thus lowering overall organisational risks in 6G deployments [68].

3.4 Challenges of Implementing Zero Trust in 6G

Despite its promise, implementing Zero Trust in 6G poses significant challenges. These range from technical barriers such as latency and scalability to operational and economic constraints. Understanding these challenges is essential for setting realistic deployment goals and for designing solutions that are both effective and practical in future 6G ecosystems.

Complexity

Adopting zero trust at the scale and dynamism of 6G introduces significant complexity in policy definition, enforcement, and auditing across heterogeneous domains [69].

Performance

Continuous verification and inline inspection mechanisms can introduce latency. Ensuring these mechanisms do not degrade 6G's low-latency guarantees is a non-trivial challenge [70].

Scalability: 6G will comprise billions of devices. Scaling zero trust policies to manage such a volume of identities and sessions without introducing bottlenecks demands innovative, decentralised solutions [71].

Cost

The infrastructure and operational changes required for ZTA—including AI engines, secure identity provisioning, and orchestration—impose substantial capital and operational costs [72].

3.5 Outlook and Future Work

As the development of 6G accelerates, further research is needed to formalise and evaluate Zero Trust frameworks tailored to its unique architectural characteristics. This includes understanding how ZTA integrates with technologies such as Reconfigurable Intelligent Surfaces (RIS), terahertz communication, and quantum-safe cryptography. Additionally, there is a growing need to benchmark the performance, cost, and scalability of ZTA components across simulated and real-world 6G deployments. Given that much of the current literature only touches on individual aspects of Zero Trust, a more holistic, system-level perspective remains a critical gap to address. Collaborative efforts between academia, industry, and standards bodies will be essential to realize a secure and trustworthy 6G ecosystem.

4. Intent-Based Networking

An intent is defined as "the formal specification of all expectations including requirements, goals, and constraints given to a technical system" [73]. Intents are therefore high-level goals provided to the system by the user. Intents at the basic hold information regarding what the system is expected or intended to perform. However, it is then up to the system to interpret, how to achieve or perform the intended task. This approach further simplifies the control surface exposed to the user compared to policy-based automation systems. The nature of policies requires complex and time-consuming efforts spent during design time to shape the system considering a variety of cases that could arise during run-time. This leads to a system that is not capable of evolving further. With increasing complexity of the 6G system, involving diverse and heterogeneous systems inter-working with each other, it becomes challenging to further scale policy-based automation. To this end, proponents suggest AI to augment policy-based automation with autonomy such that the system could learn and reach an optimal state during the run-time. The optimal state depends on the ecosystem where the system operates, for example, the manufacturing industry may need low latency service for robot operation, the process industry, however, might need highly reliable networks. The emergence of campus networks has brought forth further challenges to the 6G system, whereby operators may not have enough operational expertise or resources required to administrate the system. With the simplified control surface offered by intents, operators could just specify the goals in the form of expectations, trusting the management system takes the needed actions. Trustworthiness in the context of IBN plays an instrumental role, aligning the IBN management system with the operational framework expected by the operator. Because intents provided as high-level goals must first

be interpreted followed by translated into system actions, AI models capable of reasoning are preferred to carry out the interpretation. Due to the heterogeneous nature of telecommunication networks, although standardised, multiple data streams carry data in distinct formats exist. Reasoning models rely on an intermediate representation such as natural language in case of Large Language Model (LLM) to apply reasoning. The distinct data streams must be either fed to a use-case specific model or must be lifted into an intermediate representation.

Trustworthiness in the case of IBN therefore relies on AI to satisfy multiple criteria. When considering the intent interface, IBN has to interactively support the user by giving feedback about the status of the intended task at the level of abstraction comprehensible by the user. Furthermore, due to the existence of multiple user types, ranging from application user, service-level experts to domain-level experts, all categories of user-bases have to be equally to considered. Due to the heterogeneity, IBN has to also verify a lossless translation of distinct data streams ingested into the model to come up with right predictions. Furthermore, because there IBN can be modular based on the domain and service bifurcations, collaboration between the sub-systems is expected. With the application of AI into the networks for various purposes including automation, autonomy and management to name a few, mandates a trustworthy and explainable system. The inherent nature of AI models operating in a learned space requires mutual translation to let the user or operator know the status of a request, a task or in general the state of the system. Explainability becomes key aspect to address trustworthiness as AI agents take over automation of networks and consecutively move the network towards autonomy. As a consequence, trustworthiness evolves based on the usage of AI models.

Although intent-based networking and intent-based security can greatly easy the security orchestration and therefore the secure configuration and operation of the 6G system it also adds an additional layer of complexity and provides new attack surfaces. It is therefore essential to design the intent-based functionality with trustworthiness in mind. This e.g., covers aspects like trustworthy AI, and trustworthiness evaluation regarding the input data.

5. Confidential Computing & Remote Attestation

Confidential computing allows the secure execution of software without the need to trust the owner of the hardware. One typical application domain is cloud computing: in this case cloud computing will allow the secure execution of software even in the cloud operator is untrustworthy. There exist two different approaches to confidential computing: one is based primarily on cryptography and utilised techniques such as homomorphic encryption or secure multi-party computation. The other is primarily rooted in hardware-based security features which form a so-called Trusted Execution Environment (TEE). Such TEEs provide an isolated execution environment, which offers strong protection against other software running on the machine (including the operating system or hypervisor). This is achieved by keeping all data and software encrypted and integrity protected all time (e.g. if stored in memory or on disk) but only decrypt it inside the CPU. There exist different implementations of TEEs from different CPU manufactures such as Intel SGX, Intel TDX, AMD SEV-SNP, ARM TrustZone, NVIDIA Confidential Computing etc. One of the main differences is the level of granularity of the provided TEE. One option is, that a single TEE comprises a whole virtual machine (called confidential virtual machine, cVM). Another option is, that the TEE only comprises (parts of) a single process (called enclave). The advantage of the cVM approach is the easy deployment. There is no need to adapt existing software or learning a new programming paradigm or API. The downside of cVMs is the larger Trusted Computing Base since any component running inside the cVM (operating system etc.) can compromise the security. In case of enclaves the TCB is reduce to just the application code. The disadvantage of the enclave approach is the need of application software changes to make use of the special Enclave APIs. There exist certain frameworks which try to provide a trad-off between the need for application software adaption and the size of the TCB. In essence these frameworks provide an emulation of the operating system APIs allowing the execution of unmodified application software. The cryptographic operations involved in providing the strong security guarantees of TEEs induce certain overhead. This overhead depends e.g. on the specific TEE technology used and the application workload. As a rule of thumb, one could expect a performance penalty of 10–15%.

Remote attestation is a technique which complements TEE-based confidential computing. Remote attestation allows a (remote) verifier/relying party to check whether a given application/software is indeed running inside a TEE. Moreover, it can be verified that the software running in the TEE is not manipulated. Although confidential computing in combination with remote attestation are nice tools which supports the trustworthy execution of software without the need to trust the cloud operator there are still some challenges involved from a practical point of view. One set of challenges is related to specifying and attesting the expected status (in terms of software components) of a complex system. This becomes even more challenging, if the different components of the system changes quickly e.g. due to software updates. Another challenge is, that many confidential computing offers e.g. by the large hyper scales do not adhere to the “no need to trust the cloud provider” assumption. This is due to the fact, that they do not provide “vanilla” TEEs but make own proprietary software components part of the TEE. One needs to trust these software components and therefore ultimately the cloud provider—which undermines the concept of confidential computing.

6. Formal Methods and Formal Verification

Mobile networks are quite complex systems. So far, the complexity increased with every new generation. Although one design goal of 6G is to reduce complexity the introduction of new features such as Joint Communication and Sensing, Intent-based networking or the usage of

Artificial Intelligence in general will let the 6G-system remain a very complex system. This complexity increases the likelihood of mistakes and errors during specification and implementation of the 6G-system. Such errors can have a huge impact on the trustworthiness of the overall system. Common approaches of addressing this in terms of tests and certification are valuable but can never proof the absence of errors or contradictions.

The overall situation becomes even more challenging, since mobile networks are not only very complex, but the specifications are mainly written in natural language which opens the door for misunderstandings, misinterpretations or ambiguities. Applying formal methods and formal analysis and verification could be one way to address these challenges. Based on formal specifications formal methods could proof the absence of e.g. contradictions or security violations. Moreover, formal specifications could be the foundation for automatically derived implementations lowering the burden regarding conformance and interoperability test. Although formal verification and formal standards will be highly beneficial, there exist currently some burden which hinders it widely adoption. One of the challenges is the effort regarding the formal verification itself.

Although new methods and (semi) automated tools which were developed recently lower the effort significantly, the overall effort especially regarding the formal verification of larger systems can still be considered to be very high. Another obstacle comes with the difficulty to provide correct formal specifications—especially for non-experts in the domain of formal methods. Similarly, to the verification process itself there is ongoing effort to make the specification languages more accessible for non-domain experts. Yet it is still challenging for an average engineer to create or even easily understand formal specifications.

On potential solution could be, to provide different views regarding a given specification, e.g., a natural language one which is automatically generated from the formal specification. Such transformation processes could be supported by generative AI tools. Another challenge would be the transformation of the currently existing standard documents into a more formal representation. Although AI might be of help here as well, it is currently much more unclear to which extend this can really be done. But given the overall progress in the last years in the domain of natural language processing there is at least some hope that in the not so far future automatic conversion is doable to an extend which keeps the effort of a potentially necessary human post-processing at a manageable level.

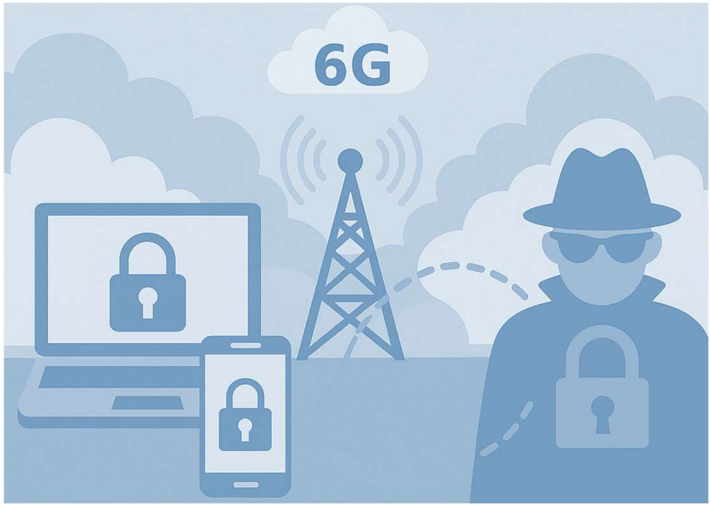
To summarize: methodologies and procedures to enable the creation of standards which make formal reasoning regarding trustworthiness easier would be helpful to support a higher level of confidence and assurance regarding the trustworthiness of the next generation mobile networks. This can be ground on scientific findings and related approaches of the past and aligned e.g., with recent activities of the IRTF research group “Usable Formal Methods”2.

7. Physical Layer Security

Physical Layer Security (PhySec) leverages the inherent physical characteristics of the communication medium, such as channel randomness, reciprocity, and spatial decorrelation, to provide security and privacy guarantees that are difficult to achieve using traditional cryptographic methods. These techniques enable the design of secure systems that offer lightweight and low-latency protection, particularly suitable for the resource-constrained and dynamic environments anticipated in 6G wireless networks. Core mechanisms include channel-based secret key generation, physical-layer authentication, secure transmission through beamforming or artificial noise injection, and anti-jamming strategies. These approaches are attractive due to their efficiency and their ability to function without the need for complex key management infrastructures or computational assumptions.

7.1 Physical Layer Security as a Key Technical Enabler for Trustworthiness

The increasing demand for trust in 6G systems, which encompasses not only security and privacy but also availability, reliability, integrity, and resilience, positions Physical Layer Security as a fundamental enabler. As elaborated in Section 2.2.2, these properties collectively define the Trustworthiness Characteristics (TCs) of a communication system. PhySec directly supports many of these characteristics. For example, it enhances confidentiality and security by mitigating spoofing and eavesdropping threats, enables device integrity through authentication based on physical-layer fingerprints, and supports availability through anti-jamming techniques. Additionally, adaptive channel estimation and re-keying mechanisms improve reliability and resilience, particularly in dynamic or adversarial environments. By operating at the lowest layer of the communication stack, PhySec forms a foundational element for trust-by-design in wireless architectures.



7.2 Addressing 6G-Specific Threats Using Physical Layer Security

Emerging 6G paradigms such as integrated sensing and communication, non-terrestrial networks including satellite and aerial platforms, and large-scale deployments of the Internet of Things significantly expand the wireless attack surface. Traditional cryptographic techniques may struggle to scale in these environments, particularly under the real-time constraints and limited computational capacity of edge devices. PhySec is well suited to such conditions due to its low overhead and reliance on the physical properties of the wireless medium. For instance, in ICAS scenarios, secure waveform design [41] combined with channel randomness to obscure radar parameters [74] reduces the likelihood that passive adversaries can extract either communication messages or sensing data. Furthermore, echo authentication based on physical-layer fingerprints helps to ensure the integrity of sensor measurements, mitigating the risk of replay or injection attacks. These capabilities are particularly relevant to the emerging Quality of Sensing (QoS) metric, which is increasingly viewed as a critical indicator of trust in applications such as autonomous systems, digital twins, and precision healthcare.

7.3 Quantum Resilience and Information-Theoretic Security

PhySec offers resistance to adversaries equipped with quantum computing capabilities, as its mechanisms are grounded in information-theoretic principles rather than computational assumptions. Classical public-key cryptographic methods, including those based on integer factorization or discrete logarithms, are vulnerable to quantum algorithms such as Shor's. In contrast, techniques such as physical-layer key generation and secure modulation schemes do not rely on computational hardness and remain secure in the presence of quantum-enabled adversaries. As demonstrated in [75] [76], symmetric cryptographic keys can be independently derived at both ends of a communication link by exploiting the reciprocity of the wireless channel. Provided that the propagation environment contains sufficient multipath richness, an eavesdropper located more than half a wavelength away typically observes an uncorrelated channel. However, this assumption is environment-dependent and may not hold in line-of-sight or low-scattering conditions.

7.4 Limitations and Open Challenges in 6G Environments

Despite its advantages, Physical Layer Security faces several limitations in practical deployment. Its performance depends strongly on the presence of channel randomness and reciprocity, which may degrade in static or line-of-sight environments [77]. Moreover, non-idealities such as hardware variation, synchronization mismatches, and limited entropy in short-range communications can reduce the effectiveness of key generation and authentication mechanisms. Scaling these techniques to large multi-user networks and integrating them with existing cryptographic protocols are active research challenges. These limitations highlight the need for hybrid security architectures that combine PhySec with higher-layer techniques, including conventional cryptography, remote attestation, and zero-trust identity frameworks.

Additional challenges emerge in specific beyond-5G contexts. For example, accurate and practical channel models are essential for high-frequency bands such as terahertz and visible light communications. Although directionality in terahertz systems improves confidentiality, the potential for eavesdropping remains and requires further study. The design of inherently secure waveforms offers another promising direction for providing confidentiality and enabling key generation with minimal overhead. In the context of aerial and satellite networks, PhySec can strengthen communication links and mitigate threats posed by malicious unmanned aerial vehicles through techniques such as three-dimensional beamforming and aerial jamming. Machine learning methods are increasingly being proposed to enhance physical-layer authentication by enabling adaptive and context-aware decision-making. However, challenges such as, computational latency, distributed learning efficiency, and device heterogeneity must be addressed to enable practical and scalable deployment.

In summary, PhySec presents an efficient and flexible set of tools for addressing emerging security threats and broader trustworthiness requirements in 6G networks. By contributing to key Trustworthiness Characteristics, including confidentiality, integrity, availability, reliability, and resilience, it supports end-to-end trust that originates at the physical layer. Its resilience against quantum attacks and suitability for real-time, resource-constrained environments make it an essential component of future wireless systems, particularly in applications requiring critical sensing, ultra-reliable low-latency communication, and infrastructure-free operation.

8. Enhanced Trustworthiness via Hardware-Based Separation

The realisation of the 6g system will require the integration of new hardware components and accelerators. This covers e.g., network accelerators in terms of smart NICs, crypto accelerators and especially AI accelerators. Some of these new hardware components will become part of more efficient and powerful UEs while others will be part of the 6G RAN and 6G core. Thereby these hardware components might come for a variety of different vendors − e.g., in the form of IP cores to be used in MPSoC designs or as chiplets to be directly integrated into the overall SoC packages. To mitigate supply chain risks which arise from potentially untrustworthy suppliers or manipulated components and component designs, the overall hardware architecture needs to support the trustworthy integration of potentially untrustworthy components.

One potential approach to achieve this and to mitigate the related supply chain risks is described in [78] and also mentioned in the Hexa-X-II

deliverables on enabling technologies for 6G devices. The basic idea of the described approach is to add separation directly at the Network-on-Chip (NoC) level. Therefore, between each hardware component connected to the NoC and the NoC itself a so-called Trusted Communication Unit (TCU) is placed. This TCU can be imagined as a kind of “firewall” controlling the access of the different hardware components (called “tiles”) to the NoC. This also covers standard hardware components like compute cores or memory (DRAM). These components all resit in their own, separate tiles. Thereby the initial state of the overall system is “no access granted”. Only the micro kernel-based operating system running on a dedicated tile can reconfigure the TCUs to allow communication between the different tiles based on well-defined access control rules. With the help of such a design a potentially malicious hardware component (e.g., an AI accelerator) will not have full NoC access, i.e. full hardware access but the components such a malicious component can influence are limited to the tiles it can communicate with.

9. Post-quantum cryptography

The 6G system needs to be post-quantum secure. Therefore, the currently ongoing activities e.g., from 3GPP or GSMA must be intensified to ensure that not only the transmission on the control plan and user plane are post-quantum secure but the whole data processing within the 6g system is post-quantum secure. This includes also data at reset and the data handling related to supporting protocols and components.

More specifically, all cryptographic operations involving symmetric cryptography have to be extended to use at least 256-bit long secret keys. According to a 3GPP study “on the support of 256-bit algorithms for 5G” [TR33.841] this transition should be comparable simple, since many protocols and message formats are already prepared for an increased key size. Nevertheless, any adaption of the protocols for the 6G system towards the support of 256-bit long keys should enable the possibility to increase the key size even further. This would make the 6G system ready to react upon threats discovered in the future since improved attack algorithms might be able to make even better use of quantum computers. This in turn could enable practical doable attacks on 256-bit symmetric cryptographic.The more challenging part is related to the transition of the asymmetric cryptographic algorithms towards post-quantum secure ones. One challenge is related to selecting a suitable set of algorithms considering the computational overhead as well as the communication overhead induced by post-quantum cryptography. Moreover, although research on post-quantum cryptography last already for many decades, some of the recently selected NIST-approved algorithms are less analysed compared to their non-post-quantum secure counterparts currently in use. Therefore, there is a non-negligible probability that cryptographic weaknesses might be discovered in these post-quantum secure algorithms. Therefore, besides the transition to post-quantum secure asymmetric cryptographic algorithms itself, this transition must happen in a way which supports great flexibility regarding the selection and deployment of the specific algorithms used. This does not only translate to flexibility regarding protocols and software components – but even hardware in case hardware security modules or crypto accelerators are used for improved security or efficiency.

Chapter 5:

Building Trustworthiness in 6G

5

Building Trustworthiness in 6G

1. Trustworthiness-by-Design principles

6G Trustworthiness by Design principles refer to foundational concepts aimed at embedding trust-enabling features into the core architecture of the network. Even if trustworthiness may be offered as a service, its features must be designed and implemented at the foundational level of 6G. Here are the core principles of Trustworthiness by Design for 6G:

- Security by Design including strong authentication mechanisms, real-time anomalies/Threats detection methods, Zero-Trust architectures, etc.
- Privacy by design using techniques like edge-based data processing.
- Resilience by design through failover mechanisms and AI-driven self-healing networks.
- Accountability and Governance by defining clearly responsibilities and traceability of decisions for all stakeholders.
- Ethics meaning that the design must avoid biases, discrimination, or unequal access.

2. Trustworthiness-by-Design for JCAS

The evolution toward 6G networks anticipates the integration of Joint Communication and Sensing (JCAS) as a cornerstone technology. This will enable novel applications by allowing the network to “communicate” as well as “sense” the physical environment. However, the collection and processing of sensing data, which may contain Personally Identifiable Information (PII), introduce significant security and privacy challenges [79]. To address these, a trustworthy JCAS architecture is essential. Figure X illustrates such an architecture, integrating security and privacy functions into both the 6G core network and the sensing units (gNBs and UEs).

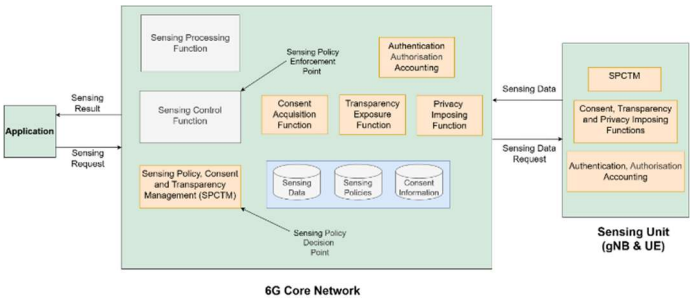


Figure 4. Trustworthiness JCAS architecture for 6G

Emerging JCAS architectures incorporate 6G core network functions such as the Sensing Control Function (SCF) and the Sensing Processing Function (SPF) for sensing management [80]. The SCF serves as the orchestrator for all sensing-related activities. When an application sends a sensing request, the SCF receives it and coordinates with other network functions to assess the feasibility of the request and configure the sensing units, such as gNBs and UEs, to collect the required data from a specific target area. The SPF acts as the data processing engine: it receives the sensing data from the sensing units, processes this information, and delivers the final sensing result to the application in the desired format. For instance, an application like traffic management system could request the number of vehicles on a crossroad.

In addition to SCF and SPF, handling potentially sensitive sensing data in JCAS systems necessitates a "trustworthy by design" approach. This can be achieved through dedicated functions that ensure compliance with data protection regulations, such as the General Data Protection Regulation (GDPR). In addition to security controls that protect the confidentiality, integrity, and availability of sensing data and services, the Authentication, Authorisation, and Accounting (AAA) functionalities are extended to support JCAS services. Every sensing request from an application must be authenticated to verify its origin and legitimacy. The system must authorise the requesting application's access to sensing results from a specific area or regarding certain subjects, based on predefined policies. The involvement of network functions and their roles in each sensing session should also be considered for accounting purposes. The architecture introduces a comprehensive framework for managing privacy, directly addressing principles that ensure compliance with data protection requirements.

The central function Sensing Policy Consent and Transparency Management (SPCTM) [79] [81] serves as the policy-keeper for all sensing operations and acts as the sensing policy decision point. It maintains up-to-date information on:

Sensing Policies

Rules that govern sensing activities, e.g., how sensing data to be collected, which entities should be involved in sensing, and which locations are prohibited from sensing.

Consent Information

Manages and verifies consent for sensing. This directly supports the privacy requirement for clear and affirmative consent from data subjects before their data can be processed. The SPCTM stores records such as who has consented, to what, for what purpose, and for how long.

Transparency Requirements

Ensures that information about sensing activities is made available, in alignment with the principle of transparency.

When a sensing request is initiated, the SCF queries the SPCTM for policies and consent information relevant to the users and geographic area specified in the request. The SPCTM, with assistance from Unified Data Management (UDM), retrieves the applicable consent records and transparency requirements, and privacy constraints like the maximum allowable data resolution to prevent over-collection. The SCF, acting as the sensing policy enforcement point, engages a set of specialized

functions to actively enforce the policy requirements defined by the SPCTM. These specialized functions include:

Consent Acquisition Function

If valid consent is not already on record, this function is triggered to obtain it from the relevant data subjects, where possible, before any data is collected.

Transparency Exposure Function

Where applicable, this function provides clear and accessible information to individuals about what data is being collected, why it is being collected, and how it will be used. This fulfills the "right to be informed" requirement of the GDPR.

Privacy Imposing Function

A critical enforcement mechanism, this function supports both the SPF and SCF in applying privacy-preserving techniques. These include not only post-processing methods (e.g., anonymisation or aggregation), but also pre-collection configurations. For example, it can instruct the SCF to configure sensing units to use lower resolutions or to avoid capturing certain types of data altogether—thereby preventing the collection of overly sensitive information from the outset, in line with the principle of purpose limitation.

In a distributed sensing environment, data collection and in some cases, processing can occur directly at gNBs and UEs. To ensure the trustworthiness of the JCAS ecosystem, it is critical to embed security and privacy capabilities within these sensing units. As shown in the Figure 4, the SPCTM and other security and privacy functions are also considered at the sensing units. A lightweight version of the SPCTM can reside on the sensing unit to manage local policies and consent. Consent, transparency, and privacy-imposing functions are also integrated to ensure that the device itself adheres to established privacy rules during data acquisition and handling. Additionally, Authentication, Authorisation, and Accounting (AAA) mechanisms must be implemented on the sensing units to prevent unauthorised use of sensing capabilities and to maintain records of their involvement in sensing activities. By distributing security and privacy functions, the JCAS architecture ensures that protections are not only managed centrally but also enforced at the point of data collection, thereby creating a more trustworthy 6G system.

3. Role of AI in ensuring Trustworthy-by-Design Networks

6G systems are envisioned as "AI-native," meaning they fundamentally integrate AI to achieve ubiquitous intelligence and support demanding applications, such as virtual reality and autonomous driving [20]. Trustworthiness in 6G is defined as the "demonstrable likelihood that the system performs according to designed behavior under any set of conditions," encompassing security, privacy, reliability, resilience, and safety. This "trustworthy-by-design" principal mandates embedding these characteristics throughout the entire 6G lifecycle [82]. AI plays a crucial role in fostering trustworthiness through several key mechanisms:

Intelligent Network Orchestration

AI-driven architectures utilize large-scale AI models for autonomous and self-optimizing network management. This approach reduces human intervention and simplifies the management of complex systems.

Enhanced Security and Resilience

AI facilitates proactive threat detection and anomaly identification using techniques such as ensemble learning, moving beyond merely reactive measures. Additionally, AI drives self-healing networks that automatically detect and correct faults, ensuring continuous operation and increased resilience.

Ensuring Privacy and Data Protection

Federated Learning (FL) enables decentralized model training, where raw user data remains on local devices while only sharing model updates. This significantly reduces privacy risks. Furthermore, Explainable AI (XAI) enhances transparency by providing understandable explanations for AI decisions, which fosters trust and accountability, especially in critical applications.

Optimizing Resource Management

AI algorithms optimize the allocation of resources, such as bandwidth and power, to improve efficiency and ensure consistent, high-quality service delivery, which is vital for reliability. Despite the benefits of AI, several challenges need to be addressed. These include managing the complexity and high computational demands of AI, as well as addressing ethical concerns such as bias and algorithmic uncertainty. Additionally, there is a need to mitigate sophisticated adversarial AI attacks.

3.1 AI as a Core Enabler for 6G Trustworthiness

The foundational integration of Artificial Intelligence within 6G networks is poised to revolutionize their design, operation, and management, making AI a core enabler for achieving the ambitious goals of trustworthiness. 6G networks are envisioned as inherently AI-native, meaning AI and Machine Learning are ubiquitous and deeply embedded across various network layers, from initial design to continuous operational execution [83]. AI-native networks transition from static, rule-based models to adaptive, learning-driven approaches, leveraging large language models (LLMs) to intelligently manage complex network operations. This deep integration facilitates innovative networking models, such as intent-based networking, which streamlines network configuration based on high-level user needs [84]. This "AI-native" approach fundamentally alters network management, shifting it from a human-centric, reactive process to an autonomous, proactive, and self-optimizing system. Such a high degree of automation reduces direct human intervention and places significant reliance on AI for executing critical, real-time decisions. Consequently, a critical requirement emerges for the network to be considered reliable. The AI systems it employs must be designed with trustworthiness as an inherent characteristic. The overall trustworthiness of the network, therefore,

becomes directly dependent on the engineered robustness, transparency, fairness, and ethical alignment of its underlying AI components [85].

Enhancing Security and Resilience through AI

AI, Machine Learning, and Deep Learning are crucial for enabling proactive threat detection and developing sophisticated mitigation strategies, ensuring 6G networks are self-sustaining. Advanced anomaly detection systems for 6G networks often employ ensemble learning (EL) techniques to identify intrusions through binary and multi-class classification [52]. AI-powered security solutions, including Intrusion Detection Systems (IDS) and anomaly detection algorithms, are vital for rapidly identifying potential threats and enabling real-time responses. For instance, the Det-RAN framework employs an AI-based design for real-time attack detection in 5G Open RAN environments [86]. AI's ability to analyze massive, dynamic datasets in real-time facilitates a crucial shift from traditional reactive security to predictive, adaptive threat intelligence, as traditional methods are insufficient for sophisticated 6G threats.

AI is instrumental in developing self-healing networks, which automatically detect network faults and failures and implement corrective actions to mitigate service degradation and minimize downtime. This automation is essential for maintaining the continuous operation of complex 6G infrastructures. The self-healing process involves AI-driven fault detection, diagnosis, and autonomous recovery actions. AI-driven systems can dynamically allocate critical resources, such as bandwidth and power, based on real-time network demands, thereby enhancing performance and user experience. Specific AI-driven approaches, such as Deep Reinforcement Learning (DRL) and Soft Actor-Critic (SAC) methods, determine optimal resource allocation policies that balance conflicting performance indicators [23] [87]. The resilience of 6G networks is significantly powered by AI models, enabling enhanced observability and rapid adaptability.

Ensuring Privacy and Data Protection with AI

Federated Learning (FL) emerges as a promising paradigm for privacy-preserving AI in 6G networks, facilitating decentralized model training without requiring the sharing of raw user data. Data remains secure on local devices, and only model updates are transmitted to a central server, significantly mitigating data breaches and privacy violations common in traditional centralized AI models. This is crucial for sensitive data from critical 6G applications, such as healthcare and autonomous vehicles. Key mechanisms include decentralized training, transmission of only model updates, secure aggregation protocols, and differential privacy [88]. Explainable AI (XAI) plays a crucial role in enhancing transparency within AI-driven decision-making processes in complex 6G-IoT environments, bridging the gap between complex machine learning outputs and human understanding. XAI fosters greater trust by providing detailed, intuitive explanations for model predictions, making AI decisions comprehensible even to non-technical stakeholders.

Optimizing Resource Management and Network Performance via AI

Artificial Intelligence is deeply embedded across various network layers to enable optimized resource allocation, improved operational efficiency, and enhanced system robust performance, particularly in intricate and dynamic 6G environments. This pervasive integration is key to managing the unprecedented demands of 6G. AI algorithms are strategically employed by both base stations and user devices for intelligent resource allocation, aiming to reduce energy consumption and significantly improve overall resource utilization [30]. AI-driven approaches, such as Deep Reinforcement Learning (DRL) methods, are being developed to determine optimal resource allocation policies that balance conflicting performance indicators like achievable data rate, overhead, and complexity. AI's unparalleled ability to process and analyse vast amounts of network data enables advanced functionalities such as predictive maintenance, accurate traffic forecasting, and proactive adjustments in network configurations [89].

Table 2. *AI's Contributions for five Trustworthiness Characteristics*

Pillar	AI's Contribution
Security	Proactive threat detection, anomaly identification, intelligent mitigation strategies, real-time response, and adversarial attack mitigation.
Privacy	Federated Learning for decentralized data training, differential privacy, secure aggregation, privacy risk management, ethical AI frameworks.
Reliability	AI-driven resource allocation, predictive maintenance, self-healing networks, dynamic network optimization, fault detection/diagnosis/recovery.
Resilience	AI-powered adaptability to circumstances, self-healing mechanisms, intelligent threat prediction, robust AI models against adversarial attacks.
Safety	AI for proactive risk analysis, intelligent control, and robust system operation to prevent critical failures.

3.2 Challenges and Critical Considerations

While AI offers immense opportunities for trustworthy 6G, its integration introduces complex challenges.

Managing AI-Driven System Complexity and Computational Demands

The deep integration of AI into 6G networks introduces unprecedented complexity, making purely human-only operations virtually impossible. AI models require significant computational power for real-time decision-making and continuous learning, which can lead to inefficiencies and bottlenecks if not appropriately managed [89]. The high data rates and intricate network topologies of 6G pose considerable challenges for AI-enabled learning and training. This necessitates new, sophisticated AI-driven management paradigms to ensure the overall trustworthiness of the communication system.

Addressing Ethical Implications, Bias, and Algorithmic Uncertainty

The ethical and societal implications of AI in 6G are paramount for maintaining public trust and ensuring regulatory compliance. Concerns exist regarding data privacy, the ethical use of AI, biased algorithms, and the misuse of personal data, particularly in sensitive sectors. The "black-box" nature of many advanced AI models complicates transparency and accountability, making it challenging to assess compliance with data protection principles and ethical guidelines [90]. The inherent uncertainty of AI algorithms in dynamic network environments raises concerns about consistent reliability and predictability. Trustworthiness in 6G extends beyond technical performance to encompass societal acceptance and ethical principles, requiring proactive development of ethical AI frameworks, Explainable AI (XAI), and robust data governance policies.

Mitigating Adversarial AI Attacks and Evolving Cyber Threats

The integration of AI into 6G significantly increases the potential for misuse or manipulation, leading to unforeseen security breaches. AI systems are susceptible to unique attacks, such as adversarial manipulations, data poisoning, and model poisoning, which target the integrity and reliability of the AI models themselves. The expansion of 6G to heterogeneous environments, including non-terrestrial networks, and reliance on multivendor components and IoT devices collectively expand the attack surface and introduce novel security challenges, such as the physical tampering of lightweight edge devices [91].

3.3 Conclusion & recommendations

AI is recognized as the foundational intelligence driving Sixth Generation (6G) systems, making its trustworthiness paramount for the successful deployment and widespread adoption of these systems. AI significantly enhances network security through proactive threat detection and self-healing mechanisms, ensures robust privacy via federated learning paradigms, improves overall reliability through intelligent resource management, and fosters transparency with explainable AI. These multifaceted contributions are indispensable for meeting the stringent demands of future applications. However, the integration of AI concurrently introduces notable challenges, including the management of inherent complexity and substantial computational demands, the imperative to address ethical implications such as algorithmic bias, and the critical need to mitigate sophisticated adversarial AI attacks. Consequently, global research and standardization efforts are actively engaged in defining comprehensive frameworks designed to embed trustworthiness throughout the entire 6G lifecycle. To advance the development of trustworthy-by-design 6G networks, it is imperative to prioritize research focused on developing AI models that are inherently robust, explainable, and ethically aligned, thereby incorporating trust and societal impact as core design criteria. Furthermore, accelerating standardization efforts for AI trustworthiness is crucial to harmonizing global standards and frameworks, ensuring interoperability and a common understanding across diverse stakeholders. Substantial investment in advanced testbeds and digital twins is also essential, as these platforms facilitate the safe development, rigorous validation, and continuous improvement of AI-driven solutions for trustworthiness. Implementing comprehensive AI lifecycle management through robust MLOps frameworks will ensure the continuous monitoring, validation, and updating of AI systems, thereby ensuring their effectiveness. Moreover, fostering cross-disciplinary collaboration among experts in telecommunications, AI/ML, cybersecurity, data privacy, and ethics is vital to addressing the multifaceted challenges holistically. Finally, proactively addressing ethical and regulatory gaps by developing clear policies and regulations for AI governance, data privacy, and algorithmic accountability tailored explicitly for 6G networks will foster public confidence and ensure the responsible deployment of these technologies. By diligently pursuing these recommendations, the global community can collectively advance toward building 6G networks that are not only technologically superior but also fundamentally trustworthy, thereby forming a reliable and secure foundation for ubiquitous intelligent connectivity.

4. User-Centric Focus

User-centric design has played a significant role in product development for a long time [92] by shifting focus from system-driven functionality to the needs, intentions, and usage contexts of end users. Unlike self-centered system design, which often prioritizes technical optimization over user experience, user-centric approaches explicitly integrate the human perspective into the design process [93]. While such an anthropocentric focus can risk overlooking alternative or non-human-centric paradigms of system design, it remains essential for establishing trustworthy 6G networks. This is particularly relevant when addressing the user's perception of self-sovereignty within the increasingly complex landscape of interconnected networks, autonomous agents, and intelligent services operating on user data. A key motivation for focusing on information storage and handling lies in its fundamental role in establishing trust across various 6G use cases. Some of the primary reasons for maintaining data storage independently from individual application providers are the increasing complexity introduced by Integrated Sensing and Communication (ICAS), the growing concern over data security, and the expanding potential of big data and AI-driven applications. By consolidating data storage across multiple applications, efficiency can be significantly improved, especially in scenarios where different applications and network functionalities require access to overlapping datasets. This approach represents a shift from traditional methods, where application-specific states are managed within individual databases. Instead, a more unified framework can be envisioned which not only addresses immediate trust-related challenges but also anticipates the evolving demands of future 6G networks regarding sovereignty of network participants. This concept can be seen as analogous to the deployment of Self-Sovereign Identity (SSI), by liberating users from fragmented single-application databases and the complexities of overseeing multiple access points for their data.

5. Safety Consideration

In critical systems like avionics and autonomous vehicles, safety is typically an application-level requirement determined by the possible damage that could arise from a malfunction. Strict development standards (like ISO 26262 [94]) applied to regulated hardware and software components functioning within a clearly defined, contained platform (the actual vehicle or aircraft) are necessary for its assurance. This platform includes onboard compute units and deterministic, wired communication networks like Time-Sensitive Networking (TSN) [95].

The introduction of 5G, and more profoundly 6G, introduces a fundamental shift by extending this safety-critical platform beyond the physical confines of the vehicle or device. Applications like advanced driver-assistance systems (ADAS), autonomous driving using collaborative perception, and remote robotics increasingly rely on external communication links (e.g., V2X via cellular networks) and potentially external compute resources. This extension creates significant trustworthiness challenges related to safety:

Loss of Control and Predictability

System designers lose direct control over the end-to-end communication path when it traverses public, shared 5G/6G networks. Unlike dedicated in-vehicle networks, these external links are susceptible to unpredictable delays, jitter, and packet loss caused by congestion or interference from non-critical traffic (e.g., entertainment streaming impacting a critical V2X message). The resulting unpredictability directly conflicts with the deterministic needs of many safety-critical functions.

Message-Level Guarantees vs. Packet-Level Mechanisms

While 5G introduced URLLC (Ultra-Reliable Low-Latency Communication) [96] aiming for low latency and high reliability, these guarantees are often defined and measured at the *packet* level. However, critical information often constitutes an entire *message* spanning multiple packets. Ensuring the timely and reliable delivery of the *complete message* is paramount from an application safety perspective. Prioritizing individual packets of a safety message might not prevent interleaving delays that cause the reassembled message to miss its end-to-end deadline. While current cellular standards provide established Quality of Service (QoS) for individual packets, the methods needed to guarantee that entire messages arrive intact and on time—such as prioritizing the whole message or reserving resources for all its parts—are crucial but significantly less developed [97].

The Trustworthiness Imperative for External Data

When safety-critical decisions (e.g., emergency braking) depend on information received from external sources via 6G (e.g., sensor data from other vehicles or infrastructure in collaborative perception), the trustworthiness of that data becomes paramount. The system must have verifiable confidence in the timeliness, integrity, and authenticity of externally sourced information before acting upon it. Building frameworks for establishing and managing this trust (potentially leveraging concepts like collaborative trust) is essential for enabling safe reliance on externally provided data.

Ambiguity in Responsibility and Liability

The extension of the safety platform into public networks creates ambiguity regarding responsibility. If an accident occurs due to a communication failure or delay on the external network, establishing liability becomes complex. Does responsibility lie with the vehicle manufacturer, the network operator, the infrastructure provider, or regulatory bodies? Clear frameworks addressing liability in these multi-stakeholder, extended-platform scenarios are currently lacking but are vital for deploying safety-critical applications dependent on 6G.

Applying Mixed Criticality Concepts

Automotive and avionics systems often employ mixed-criticality concepts, assigning different levels of assurance and resource priority based on function criticality. Extending these concepts effectively to the 6G domain, ensuring that high-criticality messages receive appropriate end-to-end preferential treatment (at the message level) over lower-criticality traffic across shared network resources, is a necessary step.

6. Standardisation, Harmonization & Regulatory Frameworks

Building a trustworthy network requires compliance with multiple security standards, regulations, and frameworks. Whether it's a corporate network, a financial payment system, or a decentralized blockchain network, regulatory frameworks like ISO, NIST, and GDPR ensure that data integrity, user privacy, and secure communication are maintained. The evolving landscape of AI, cloud computing also brings new standards into play, further ensuring the trustworthiness of modern networks. By adhering to these standards, organizations can ensure their networks remain secure, reliable, and trustworthy. Cybersecurity frameworks and standards are shaped by a diverse range of organizations and directives.

Globally, the International Telecommunication Union (ITU) takes a leading role with its ITU-T Recommendations and the Global Cybersecurity Index, while organizations such as the World Economic Forum and OECD contribute to building trust in digital ecosystems. In the European Union, the NIS2 Directive and GDPR establish comprehensive measures for network security and data protection. The AI Act, coming into force in August 2026, harmonizes regulations for artificial intelligence across the EU. Moreover, ENISA, the EU Agency for Cybersecurity, further supports resilience through the Cybersecurity Act framework.

When it comes to industry-specific areas, foundational standards such as ISO/IEC 27001 and ISO/IEC 27032 serve as the pillars of global information security. These are further strengthened by the innovative work of the Cloud Security Alliance, which enhances security practices within the cloud domain. Germany stands out with its robust regulatory approach, spearheaded by the Federal Office for Information Security (BSI). The IT-Sicherheitsgesetz and its updated version, IT-Sicherheitsgesetz 2.0, enforce strict cybersecurity requirements for critical infrastructure operators and digital service providers. The KRITIS framework safeguards vital sectors like energy, healthcare, and telecommunications. Additionally, the BSI Grundschutz provides a toolkit of best practices for implementing comprehensive cybersecurity measures.

At the EU level, Germany actively supports the GDPR and is implementing the NIS2 Directive to enhance security in critical sectors. Industry-specific standards also play a crucial role, including ISO/IEC 27001, widely adopted in Germany for information security management systems (ISMS). Other notable standards like VDI/VDE 2182 address industrial automation security, while DIN SPEC 27071 focuses on IT security for SMEs, reflecting Germany's commitment to supporting businesses of all sizes. Further strengthening its cybersecurity landscape, Germany enforces regulations such as the Telekommunikationsgesetz (TKG), which ensures secure telecommunications networks, and the Energiewirtschaftsgesetz (EnWG), which sets security requirements for the energy sector.

7. Test and Certification

Standardized security tests and certifications are an important building block for trustworthy networks. Testing security functionalities assures that products are implemented correctly in terms of security and that vulnerabilities or security deficits are detected and fixed prior to product use. Security tests need to be standardized in order to ensure the consistent application and comparability of the tests. Certification by independent bodies is an elementary tool to review, assess and improve proper implementation. Three parties are involved in the process of such an independent product certification. The vendor of the product to be certified has to submit an application for certification to an independent certification body. The vendor then tasks a testing facility approved for the specific certification scheme to carry out the security tests and prepare a test report. The certification body issues or refuses the certificate on the basis of the test report. This independent testing and granting of the certificate strengthen the trustworthiness of the product. Testing and certification is a proven measure to reinforce trust in products, as it is applied in various industrial sectors, e.g. the international Common Criteria scheme to certify security of network equipment or the VDE (Verband Deutscher Elektrotechniker) certification in Germany that assures safety of electrical products.

Based on GSMA NESAS (Network Equipment Security Assurance Scheme) framework, the Federal Office for Information Security (BSI) in Germany has developed a scheme for the certification of network products used in mobile networks. This national certification scheme is called NESAS Cybersecurity Certification Scheme – German Implementation (NESAS CCS-GI). BSI has also published a technical guideline which describes approved schemes for certification of critical components in mobile networks. Especially in 5G networks, it is important as there is a regulatory obligation in Germany as of January 2026 that only certified products can be used if it is categorized as a critical component. The basis of the certification with NESAS CCS-GI are the Security Assurance Specifications (SCAS) of 3GPP which define security tests for general application and specific tests for selected network component specified by 3GPP. BSI is active in the standardization of these tests in order to assess and improve mobile network security. As these tests are discussed within an international standardization organization, it can be assumed that the output of these tests increases trustworthiness within mobile networks worldwide. Specifically, certification with NESAS CCS-GI provides a robust metric to strengthen security as a trustworthiness characteristic. However, regulatory requirements additionally need to enforce that testing and certification have to be mandatory for mobile network operators. Only

then, testing and certification can contribute to an increase in trustworthiness in public mobile networks. Additionally, it is necessary to expand requirements in 6G standardization for assessing and assuring trustworthiness in future mobile networks. This will ensure that trustworthiness itself could be independently tested and certified if necessary.

8. Leveraging 6G technology for enhanced Environmental Awareness

Building trustworthiness in 6G systems requires not only securing the network and ensuring data privacy but also adopting a responsible and sustainable approach to deployment. One of the most effective ways to reinforce this trust is by leveraging 6G's capabilities to enhance environmental awareness.

As environmental concerns increasingly influence public policy and user behavior, integrating ecological intelligence into 6G infrastructure can significantly contribute to the network's perceived trustworthiness.

6G is expected to deliver advanced sensing, AI-driven decision-making, and global connectivity. Embedding real-time environmental monitoring into its architecture is therefore essential. This includes the ability to track air and water quality, monitor greenhouse gas emissions, detect deforestation, and assess ecosystem health—using distributed sensor networks and intelligent edge devices.

Moreover, increasing transparency regarding the environmental impact of 6G systems can strengthen trust among stakeholders—from regulators and businesses to the public and end-users. By providing verifiable environmental data through secure technologies such as blockchain, 6G networks can help combat misinformation and the misrepresentation of environmental performance, thereby supporting informed decision-making and regulatory compliance.

To achieve truly trustworthy 6G systems, developers and policymakers must ensure that environmental sensing technologies are deployed ethically, safeguard user privacy, and promote equitable access. With these safeguards in place, 6G can serve as a powerful enabler of environmental awareness—demonstrating its alignment with global sustainability goals and reinforcing public trust.

Environmental Sensing

With its anticipated capabilities such as terahertz communications, distributed massive MIMO and intelligent surfaces, 6G can radically enhance environmental sensing and awareness. They can be deployed to create a hyper-connected environmental system, enabling the monitoring of environmental variables across areas, the integration of autonomous systems for rapid response to environmental hazards and the context-aware adaptation of network behaviour.

Alignment with Global Sustainability and Climate Goals

Environmental awareness and responsiveness in 6G must align closely with international frameworks such as the UN Sustainable Development Goals (SDGs), particularly Goals 11 regarding Sustainable Cities, Goal 13 for Climate Action and Goal 15 for the Life on Land. Another international framework is the EU Green Deal and Digital Decade Policy Programme, which accentuate the convergence of digital transformation and environmental responsibility.

Ethical Considerations and Deployment Equity

To implement trustworthiness into the system-design, 6G-enabled environmental awareness must be deployed ethically and equitably. This includes protecting data privacy, especially in use cases where environmental sensors may capture human movement or personal information. Also, it includes ensuring technology access across geographic and economic boundaries and avoiding technological determinism.

Conclusion and Call for Actions

1. Conclusion

As we move towards the deployment of 6G systems, trustworthiness must become a foundational pillar of the system design. Its importance has reached a critical juncture, especially for countries like Germany and across the European Union.

This white paper has highlighted the critical need for trustworthiness in 6G, especially with the increasing reliance on mobile networks for critical infrastructure, emerging technologies like AI-native architectures, non-terrestrial networks, and integrated sensing and communications.

Trustworthiness in 6G is not a singular concept but a multidimensional attribute encompassing security, privacy, resilience, safety, and reliability. Trustworthiness must be designed into every layer of the 6G system — from hardware and protocols to user interactions and regulatory frameworks. This "trustworthiness-by-design" approach must be supported by measurable and verifiable mechanisms, informed by both objective metrics and stakeholder expectations.

Key enablers — including decentralized identity, AI-driven anomaly detection, confidential computing, and zero-trust architectures — will play a central role in achieving this vision. At the same time, addressing emerging risks such as AI threats, quantum-enabled attacks, and privacy vulnerabilities in sensing applications will be essential to ensuring user confidence and societal trust.

Trustworthiness in 6G must be viewed as an ecosystem-level goal that requires cooperation among all stakeholders — telecom operators, equipment manufacturers, software vendors, researchers, regulators, standardisation bodies, and end users.

2. Preliminary selection for "Call for Actions"

Below we present a preliminary list of items for a "Call for Actions". This list was generated by consulting members of the working group "Trustworthiness" of the 6G platform, Germany. It is considered preliminary since we hope to receive more feedback from other members of the working group and in general interested stakeholders. Therefore, this list will be reworked in the upcoming weeks to eventually reflect a harmonised view of the 6G platform German on the important next steps which needs to be taken to ensure that the 6G system will fulfil the trustworthiness requirements.

Standardisation & Regulation

- foster activities/research towards more formal standards
- strengthen efforts to make more security related options mandatory in the standards
- rethink existing standards and architecture with real zero trust in mind
- create regulatory base especially regarding data protection regulation for JCAS
- remove regulatory or organisational barriers related to resilience, e.g. which might hinder the setup of temporary mobile networks in case of natural disasters or large scale attacks
- reduce complexity, try to get rid of outdated technologies and the huge variety of options
- strengthen the effort to let 3GPP include trustworthiness in the 3GPP standards

Preventive Measures

- strengthen the efforts to make JCAS trustworthy and especially privacy friendly by technical means
- push for post-quantum security from Day-0

- use remote attestation among Core and RAN components to enable the 6G components to evaluate the trustworthiness status among each other (related to the point of exposure of the trustworthiness status)
- enable/use confidential computing if components are run in public clouds
- safeguard AI-related decision making; use trustworthy/XAI and allow always huma-controlled fall backs
- rethink/improve resilience by technical means
- strengthen the reliability of connections especially wrt. safety critical applications
    - enable assurance regarding real-time capabilities under restricted resources while preserving the overall trustworthiness – not only for safety critical applications but also for XR/AR.
- strengthen decentralised user/identity management – support SSI
    - One pain point of decentralization is the MNO's subscriber data, which usually is the central point of any core network architecture. To push the decentralization of 6G core networks, we need distributed subscriber data storage solutions that adhere to the MNO's security policies.
    - We regularly study approaches to distributing 6G core network functionality, particularly the control plane functions. One non-trivial issue here is establishing trust (a) between distributed network functions and (b) between users and the network, as data to confirm SIM cards might not be available during crises.
- strengthen network slice orchestration to allow timely yet sill secure/trustworthy establishment of new network slices (e.g., for emergency situations)
- strengthen the possibilities for users to control which data are collected about them and how these data are distributed/shared/processed
- Develop a 'Trustworthy-by-Design' platform to support end-to-end trustworthiness processes essential for next-generation 6G systems.

Monitoring & Transparency

- strengthen the monitoring capabilities of the whole 6G system to enhance the detectability of attacks or anomalies.
- Expose the trustworthiness status (potentially using easily perceivable “trustworthiness labels”) of the network to the users and operators enable predictive channel quality estimations like MCS degradation (this point is related to the overall monitoring improvements)

References

[1] ISO/IEC, "Trustwhorthiness - Vocabulary, Technical Specification ISO/IEC TS 5723:2022, First Edition.," Juli 2022.

[2] NIST, "Engineering Trustworthy Secure Systems," *NIST Special Publication NIST SP 800-160v1r1, [Online] https://doi.org/10.6028/NIST.SP.800-160v1r1*, November 2022.

[3] V. José María Jorquera et al, "Building Trust in the Era of 6G: A Level of Trust Assessment Function for Cloud Continuum," in *Joint European Conference on Networks and Communications & 6G Summit (EuCNC/6G Summit)*, 2024.

[4] NGMN, *6G Trustworthiness Considerations v1.0.*, October 2023.

[5] Catalina Mladin, "draft0_S1-252286_Update to 6.1 UC on 6G Infrastructure Optimization," *https://www.3gpp.org/ftp/Meetings_3GPP_SYNC/SA1/Inbox/drafts/*, 2025.

[6] ISO/IEC, *Software and systems engineering — Capabilities of software safety and security verification tools, ISO/IEC 23643:2020. Sec 3.16.*, June 2020.

[7] ISO, *Market, opinion and social research, including insights and data analytics — Vocabulary and service requirements, ISO 20252:2019, Sec 3.49.*, February 2019.

[8] ISO, *Information processing systems — Open Systems Interconnection — Basic Reference Model - Part 2: Security Architecture, ISO 7498-2:1989.*, February 1989.

[9] ISO/TS, *Health informatics — Document registry framework, ISO/TS 27790:2009.*, December 2009.

[10] ISO/TS, *Health informatics — Security and privacy requirements of EHR systems for use in conformity assessment, ISO/TS 14441:2013.*, December 2012.

[11] 3GPP, *Service requirements for the 5G system, 3GPP TS 22.261.*, August 2016.

[12] IEC, *International Electrotechnical Vocabulary (IEV) - Part 192: Dependability, IEC 60050-192.*, 2015.

[13] ETSI, *Network Functions Virtualisation (NFV); Terminology for Main Concepts in NFV, Group Report ETSI GR NFV 003.*, Decenber 2024.

[14] IFIP Working Group 10.4, "Dependable Computing and Fault Tolerance," *https://www.dependability.org/?page_id=265*, 2024.

[15] L. Khaloopour et al., "Resilience-by-Design in 6G Networks: Literature Review and Novel Enabling Concepts," *IEEE access*, vol. 12, pp. 155666–155695, 2024.

[16] ISO, "Sustainability in buildings and civil engineering works — General principles, ISO 15392:2019," *December 2019*.

[17] ISO/IEC/IEEE, "Systems and software engineering — Software life cycle processes, ISO/IEC/IEEE 12207:2017," November 2017.

[18] ISO, "Systèmes de management de la qualité — Principes essentiels et vocabulaire, ISO 9001:2015," September 2015.

[19] R. Chataut et al., "6G networks and the AI revolution—Exploring Technologies, applications, and emerging challenges," *Sensors*, vol. 24, no. 6, p. 1888, 2024.

[20] W. Yiying et al., "SIX-trust for 6G: Towards a secure and trustworthy 6G network," *arXiv preprint arXiv:2210.17291*, 2022.

[21] Mika, et al. Ylianttila, "6G white paper: Research challenges for trust, security and privacy," *arXiv preprint arXiv:2004.11665*, 2020.

[22] Weisi. Guo, "Explainable artificial intelligence for 6G: Improving trust between human and machine," *IEEE Communications Magazine* , pp. 39-45, 2020.

[23] et al. Siriwardhana Yushan, "AI and 6G security: Opportunities and challenges," *IEEE Joint European Conference on Networks and Communications & 6G Summit*, pp. 6161-621, 2021.

[24] Gerhard P and Holger Boche Fettweis, "On 6G and trustworthiness," *Communications of the ACM*, vol. 65, no. 4, pp. 48-49, 2022.

[25] Letaief Khaled B et al., "The roadmap to 6G: AI empowered wireless networks," *IEEE communications magazine*, vol. 57, no. 8, pp. 84-90, 2019.

[26] Basaran et al., "XAInomaly: Explainable, Interpretable and Trustworthy AI for xURLLC in 6G Open-RAN," *IEEE 3rd International Conference on 6G Networking*, 2024.

[27] Yulei. Wu, "Ethically responsible and trustworthy autonomous systems for 6G," *IEEE Network*, vol. 36, no. 4, pp. 126-133, 2022.

[28] M. Liyanage, "Explainable AI for B5G/6G: Technical aspects, use cases, and research challenges.," *arXiv preprint arXiv:2112.04698* , 2021.

[29] Botez R et al., "Redefining 6G Network Slicing: AI-Driven Solutions for Future Use Cases," *Electronics*, vol. 14, no. 2, p. 368.

[30] Cui Qimei et al., "Overview of AI and communication for 6G network: fundamentals, challenges, and future research opportunities," *Science China Information Sciences*, vol. 68, no. 7, p. 171301, 2025.

[31] N. Su et al., "Secure radar-communication systems with malicious targets: Integrating radar, communications and jamming functionalities," *IEEE Transactions on Wireless Communications*, vol. 20, no. 1, pp. 83–95, 2021.

[32] Z. Wei et al., "Integrated sensing and communication signals toward 5G and 6G: A survey," *IEEE Internet of Things Journal*, vol. 10, no. 13, pp. 11068–11092, 2023.

[33] J. Xu et al., "Anti-jamming design for integrated sensing and communication via aerial IRS," *IEEE Transactions on Communications*, 2024.

[34] X. Chen et al., "A survey on DDNOS attack, prevention, and mitigation techniques," *International Journal of Distributed Sensor Networks*, vol. 13, no. 12, pp. 1–26, 2017.

[35] Understanding and responding to distributed denial-of-service attacks, "Cybersecurity and I. S. A. (CISA)," 2024.

[36] Y. Ma et al., "Wifi sensing with channel state information: A survey," *ACM Computing Surveys (CSUR)*, vol. 52, no. 3, pp. 1–36, 2019.

[37] M. Kumar et al., "A survey on sybil attack detection techniques in IoT-based wireless sensor networks," *PeerJ Computer Science*, vol. 8, p. e673, 2022.

[38] J. R. Douceur, "The sybil attack," *International Workshop on Peer-to-Peer Systems*, pp. 251–260, 2002.

[39] H. Ambalkar et al., "Adversarial human activity recognition using Wi-Fi CSI," *IEEE Canadian Conference on Electrical and Computer Engineering (CCECE)*, pp. 1-5, 2021.

[40] Y. Zhou et al., "Wiadv: Practical and robust adversarial attack against WiFi-based gesture recognition system," *The Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 6, no. 2, 2022.

[41] A. K. Boroujeni et al., "Enhancing Frequency Hopping Security in ISAC Systems: A Physical Layer Security Approach," *4th International Symposium on Joint Communications & Sensing (JC&S)*, pp. 1-6, 2024.

[42] A. Dimas et al., "On Radar Privacy in Shared Spectrum Scenarios," *IEEE international Conference on Acoustics Speech and signal Processing ICASSP*, pp. 7790-7794, May 2019.

[43] Rechert et al., "Reclaiming Location Privacy in Mobile Telephony Networks—Effects and Consequences for Providers and Subscribers," *IEEE Systems Journal*, vol. 7, no. 2, pp. 211-222, 2013.

[44] N. Shenm et al., "Privacy-Preserving Location Sharing Mechanism in Mobile Online Social Networks," *9th International Conference on Broadband and Wireless Computing, Communication and Applications*, pp. 312-316, 2014.

[45] RFC, "Intent-Based Networking - Concepts and Definitions, RFC 9315," October 2022.

[46] I. Ahmad et al., "Security in Intent-Based Networking: Challenges and Solutions," *IEEE Conference on Standards for Communications and Networking (CSCN)*, pp. 296-301, 2023.

[47] F. de Trizio et al., "A Novel Malicious Intent Detection Approach in Intent-Based Enterprise Networks," *20th International Conference on Network and Service Management*, pp. 1-7, 2024.

[48] S. Reza et al., "A Survey of Self-Sovereign Identity Ecosystem," *Security and Communication Networks*, p. 8873429, 2021.

[49] F. Md Sadek et al., "In search of self-sovereign identity leveraging blockchain technology," *IEEE access 7*, pp. 103059-103079, 2019.

[50] C. Abdelaali et al., "6G for bridging the digital divide: Wireless connectivity to remote areas," *IEEE Wireless Communications*, pp. 160-168.

[51] Y. Rumesh et al., "Federated Learning for Anomaly Detection in Open RAN: Security Architecture Within a Digital Twin," *Joint European Conference on Networks and Communications & 6G Summit (EUCNC/ 6G Summit)*, pp. 877-882, 2024.

[52] M. M. Saeed et al., "Anomaly Detection in 6G Networks Using Machine Learning Methods," *Electronics*, vol. 12, no. 15, p. 3300, 2023.

[53] A. Schösser et al., "Advancing Spectrum Anomaly Detection through Digital Twins," *IEEE Communications Magazine*, pp. 1-7, 2024.

[54] A. Schösser et al., "Leveraging the Digital Twin Channel for Spectrum Anomaly Detection: An Experimental Study," *Proceedings of IEEE 5th International Symposium on Joint Communications & Sensing*, 2025.

[55] S. B. Mallikarjun et al., "Machine Learning Based Anomaly and Intrusion Detection to mitigate DoS and DDoS attacks in Private Campus Networks," *Mobilkommunikation; 28. ITG-Fachtagung* , pp. 19-24, 2024.

[56] Hexa-X II, "End-to-end system evaluation results from the interim overall 6G system," *Deliverable D2.4*, 2024.

[57] A. Kumar et al., "Malicious Lateral Movement in 5G Core With Network Slicing And Its Detection," *33rd International Telecommunication Networks and Applications Conference*, pp. 110-117, 2023.

[58] S. Rose et al., "Zero Trust Architecture," *NIST Special Publication*, pp. 800-207, 2020.

[59] M. M. Fouad et al., "Zero Trust in Network Slicing: Towards Secure 6G Architectures," *IEEE Communications Surveys & Tutorials*, vol. 24, no. 4, pp. 2710–2734, 2022.

[60] D. R. Kuhn, "Zero Trust: A Cybersecurity Paradigm Shift," *IEEE Computer Society*, vol. 54, no. 5, pp. 66–70, 2021.

[61] L. U. Khan et al., "AI-Empowered 6G: Security Challenges and Research Directions," *IEEE Network*, vol. 36, no. 1, pp. 170–177, 2022.

[62] Zhang et al., "Blockchain-Based Authentication and Authorization for 6G Networks," *IEEE Internet of Things Journal*, vol. 8, no. 9, pp. 7613–7627, 2021.

[63] T. Taleb et al., "Security Automation for Beyond-5G Networks," *IEEE Wireless Communications*, vol. 27, no. 5, pp. 118-125, 2020.

[64] Y. Liu et al., "Threat Intelligence in Zero Trust Security Models for 6G," *IEEE Transactions on Information Forensics and Security*, vol. 18, pp. 789-802, 2023.

[65] R. Singh and K. Chopra, "Mitigating Insider Threats in 6G Using ZTA," *Future Generation Computer Systems*, vol. 128, pp. 198-209, 2022.

[66] M. Bennis et al., "Toward Resilient and Trustworthy 6G," *Proceedings of the IEEE*, vol. 109, no. 10, pp. 1811-1830, 2021.

[67] Z. Zhang et al., "6G Use Cases and Enabling Technologies," *IEEE Vehicular Technology Magazine*, vol. 17, no. 1, pp. 32-44, 2022.

[68] M. Ylianttila et al., "Zero Trust Networking for 6G," *IEEE Open Journal of the Communications Society*, vol. 2, pp. 1809-1824, 2021.

[69] X. Li et al, "Zero Trust Deployment Challenges in Heterogeneous Networks," *ACM Computing Surveys*, vol. 55, no. 3, pp. 1–35, 2023.

[70] A. Gupta et al., "Latency-Aware Zero Trust in URLLC for 6G," *IEEE Transactions on Industrial Informatics*, vol. 18, no. 6, pp. 3890–3901, 2022.

[71] H. X. Nguyen et al., "Scaling Zero Trust for the Massive 6G Device Landscape," *IEEE Communications Magazine*, vol. 61, no. 4, pp. 80-86, 2023.

[72] P. Wang and X. Zhou, "Economic Impacts of Security Transformations in 6G,"," *Journal of Network and Computer Applications*, vol. 201, no. 103334, 2022.

[73] IG1253, "Intent in Autonomous Networks ," *TM Forum*, 2021.

[74] A. K. Boroujeni et al., "Frequency hopping waveform design for secure integrated sensing and communications, arXiv preprint arXiv:2504.10052 (2025)," 2025.

[75] M. Bloch and J. Barros, "Physical-Layer Security: From Information Theory to Security Engineering," *Cambridge University Press*, 2011.

[76] A. Mukherjee et al., "Principles of physical layer security in multiuser wireless networks: A survey," *IEEE Commun. Surveys Tuts.*, vol. 16, no. 3, pp. 1550–1573, 2014.

[77] Diana PM, SÁNCHEZ, José DV, et ALVES, Hirley OSORIO, "Physical-Layer Security for 5G and Beyond," *Wiley 5G Ref: The Essential 5G Reference Online*, pp. 1-19, 2019.

[78] F. Pauls et al., "Trust-minimized Integration of Third-Party Intellectual Property Cores," *20th International SoC Design Conference*, 2023.

[79] et al. P. Dass, "Addressing privacy concerns in joint communication and sensing for 6G networks: challenges and prospects," *Annual Privacy Forum*, pp. 87-111, 2024.

[80] Hexa-X-II project, "Deliverable D3.3 Initial analysis of architectural enablers and framework," 2024.

[81] P. Gersing et al., "Architecture Proposal for 6G Systems Integrating Sensing and Communication," *arXiv preprint arXiv:2411.10138* , 2024.

[82] B. Veith et al., "The road to trustworthy 6G: A survey on trust anchor technologies," *IEEE Open Journal of the Communications Society 4*, pp. 581-595, 2023.

[83] Vo Thi Kim. Anh, "The rise of AI in 6G networks: A comprehensive review of opportunities, challenges, and applications," *IEEE International Conference on Advanced Technologies for Communications ATC*, 2024.

[84] V. Deepak et al, "AI in 6G network security and management," *Reshaping CyberSecurity with Generative AI Techniques. IGI Global*, pp. 173-200, 2025.

[85] A. Nechi et al., "Practical trustworthiness model for DNN in dedicated 6G application," *19th IEEE International Conference on Wireless and Mobile Computing, Networking and Communications*, 2023.

[86] A. Scalingi et al., "Det-RAN: Data-driven cross-layer real-time attack detection in 5G open RANs," *IEEE INFOCOM 2024-IEEE Conference on Computer Communications*, 2024.

[87] D. Hirsch et al, "Efficient AI-based Attack Detection Methods for Sensitive Edge Devices and Systems," *Advancing Edge Artificial Intelligence. River Publishers*, pp. 177-196, 2024.

[88] R. Teixeira et al, "Leveraging decentralized communication for privacy-preserving federated learning in 6G Networks 233 (2025): 108072.," *Computer Communications*, vol. 233, p. 108072, 2025.

[89] Y. Sanjalawe et al., "A Review of 6G and AI Convergence: Enhancing Communication Networks With Artificial Intelligence," *IEEE Open Journal of the Communications Society*, 2025.

[90] Keivan. Navaie, "Personal Data Protection in AI-Native 6G Systems ," *arXiv preprint arXiv:2411.03368*, 2024.

[91] Kaur Navneet and Lav Gupta, "Securing the 6G–IoT Environment: A Framework for Enhancing Transparency in Artificial Intelligence Decision-Making Through Explainable Artificial Intelligence," *Sensors*, vol. 25, no. 3, p. 854, 2025.

[92] Jan, et al. Gulliksen, "Key principles for user-centred systems design," *Behaviour and Information Technology* , vol. 22, no. 6, pp. 397-409, 2003.

[93] Birger Sevaldson, "Beyond user centric design," pp. 516-525, 2018.

[94] ISO 26262, ""Road vehicles – Functional safety, Edition 2," Dec. 2018.

[95] Janos, Lucia Lo Bello, and Craig Gunther Farkas, "Time-sensitive networking standards," *IEEE Communications Standards Magazine* , vol. 2, no. 2, pp. 20-21, 2018.

[96] Petar, et al. Popovski, "Wireless access in ultra-reliable low-latency communication (URLLC)," *IEEE Transactions on Communications*, vol. 67, no. 8, pp. 5783-5801, 2019.

[97] Delia, and Pedro Merino Rico, "A survey of end-to-end solutions for reliable low-latency communications in 5G networks," *IEEE Access 8*, pp. 192808-192834, 2020.

Scan to learn more

https://www.lites.tf.fau.de/en/6g-platform/