Ericsson White Paper
GFTL-24:000856 Uen
July 2024

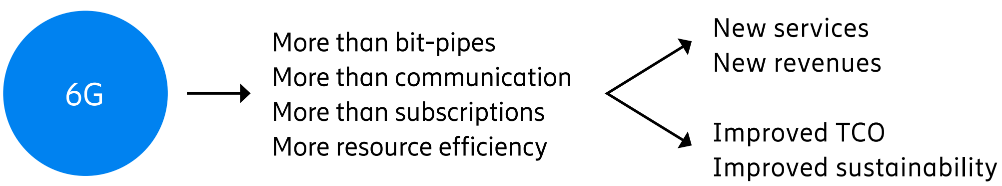
Co-creating a cyber-physical world

Content

Introduction

6G development is now entering a more concrete phase of regulation [1] and standardization [2]. Even though much has happened in the industry, the Ericsson 6G vision first announced in 2020 remains relevant for transforming 2030's business and society with 6G technology. We envision a future where individuals and digital entities have access to enhanced communication services delivered securely, efficiently, and sustainably. In response to this, many of the world's leading telecom stakeholders are converging on the problems that need to be addressed and the goals that should be accomplished (discussed in [3]), for a continued successful global collaboration in communication networks. Simultaneously, the insights will be drawn from the ongoing expansion of 5G networks worldwide.

The future networks must tackle numerous societal challenges including resilience and sustainability while also supporting new and evolved use cases and capabilities, as well as the traffic increase, predicted to be 3X in 2029 [4]. The network services must continue to respond to actual needs of people and enterprises and be delivered in a way that fits the users' needs and creates value. Ericsson's proposal for achieving this – the 6G platform – aims to deliver advanced telecommunication networks that offer enhanced optimization and efficiency for communication service providers, along with new monetization opportunities. It will also grant access to a wider set of network services to enterprises and developers, while delivering new experiences to end consumers. 6G should support applications with guaranteed services involving more than communication, with value in focus.



By leveraging technological breakthroughs and applying new design concepts, flexibility and efficiency can be achieved while meeting the rigorous demands of new services.
On the 6G platform, continued enhancements in the communication technology will be complemented with adjacent services beyond communication, providing a wider range of solutions to a larger customer base compared to 5G. Generative AI, among other AI technologies, will play a crucial role in automation of network management and operations [5]. The networks should build on and expand further on the 5G capabilities and be better integrated into the rest of the information and communications technology (ICT) ecosystem of cloud platforms and applications. Specifically, drawing from the insights gained from 5G, 6G networks will need to address challenges related to complexity, operating costs, device scale, and expansion into new business segments.

This white paper presents a scenario for future communication needs and beyond, gives an overview of the main components of the 6G network platform, and a projection of its potential applications. Finally, a roadmap is being outlined to establish viable 6G systems by the 2030s, ready to transform business and society.

Cyber-physical world of 2030

As we approach the 2030s, rapid digitalization is expected to extend into all domains, with AI-powered automation available at all levels. Digitalization brings new opportunities for learning and knowledge creation, leading to improved efficiency, business innovation, and enhanced quality of lives. This development will result in physical things, people and activities being connected through a digital domain of intelligence and data, creating a fully cyber-physical world. To this end, 6G should fulfill the role of being the critical bridge between the domains, placing communication networks in the center of future society — even more than today's networks.

Shaping 6G for this scenario involves not only enhancing the communication capabilities and adding supporting functionality, but also ensuring that networks align with the expectations of a sustainable society, providing reliable digital interactions wherever needed. A truly cyber-physical world should include everyone and still have a small footprint. To play a critical role in society, resilience and integration of 6G into other ecosystems is important. As we move towards digitalization, ensuring data privacy and secure communication becomes increasingly critical.

Today's 5G use cases, which are dominated by enhanced mobile broadband (eMBB) and fixed wireless access (FWA), will move toward improved performance and increased service differentiation and guarantees, adding value to many sectors. Simultaneously, the demand for expanded IoT and critical services will rise owing to increasing cyber-physical interactions. And therefore, new services are needed that provide rich personal experiences, relevant enterprise and application solutions, and enhancements for society in general.

6G should support several key aspects of cyber-physical interactions and related use cases:

Immersion
Digital objects seamlessly integrated with the physical environments to create experiences, where distance between people and places no longer serves as a barrier to interaction. Use cases such as **mixed reality** (MR) rely on high-capacity data links with assured quality of service for advanced head mounted devices, supported by network compute offload for better battery performance, aiding in localization and mapping, and ensuring privacy of users and bystanders.

Inclusion
Getting access to internet services that improves welfare of people and growth of businesses, without being limited by the location or activity. Wireless connectivity spans across networks to deliver **global internet**, leveraging a combination of macro cells, long range base station towers, low-Earth orbit satellites, and denser deployments, optimizing the cost through spectrum sharing and energy performance through micro-sleep.

Resilience
Ensuring critical service delivery during disruptive events, such that digital wireless solutions can be applied widely in society using **resilient connectivity**. For instance, sensors can monitor a wide area for natural disasters, surveil borders, and provide guidance for emergency services. Here, network connectivity should enable network service providers to ensure high service availability, through duplicated and fallback accesses, observability of and adaptability to circumstances, powered by AI models.

Awareness
Providing time-critical and accurate information about airborne and ground vehicles, objects and people in streets, and of the surroundings, to aid navigation and prevent collisions. For this, networks can use the radio interface to localize connected and unconnected objects and become an engine for **spatial data**, provide compute for mapping and route suggestions.

Insight
Creating webs of connected sensors and actuators in industries, cities, and homes integrated into **massive digital twins** where planning, analysis, and control can be applied. By using IoT solutions on larger (low power wide area, LPWA) and shorter distances (zero energy or ambient IoT), and applying AI and compute capabilities, networks will be the key enabler for the digitalization of any process.

Autonomy
Enabling machines to operate with greater autonomy (for example, collaborating robots and assisted cars). Networks can support **autonomous mobility** of vehicles by delivering full 3D coverage of reliable data links for time-critical interactions, as well as using spatial data and Digital Twin tools to achieve safe interworking.

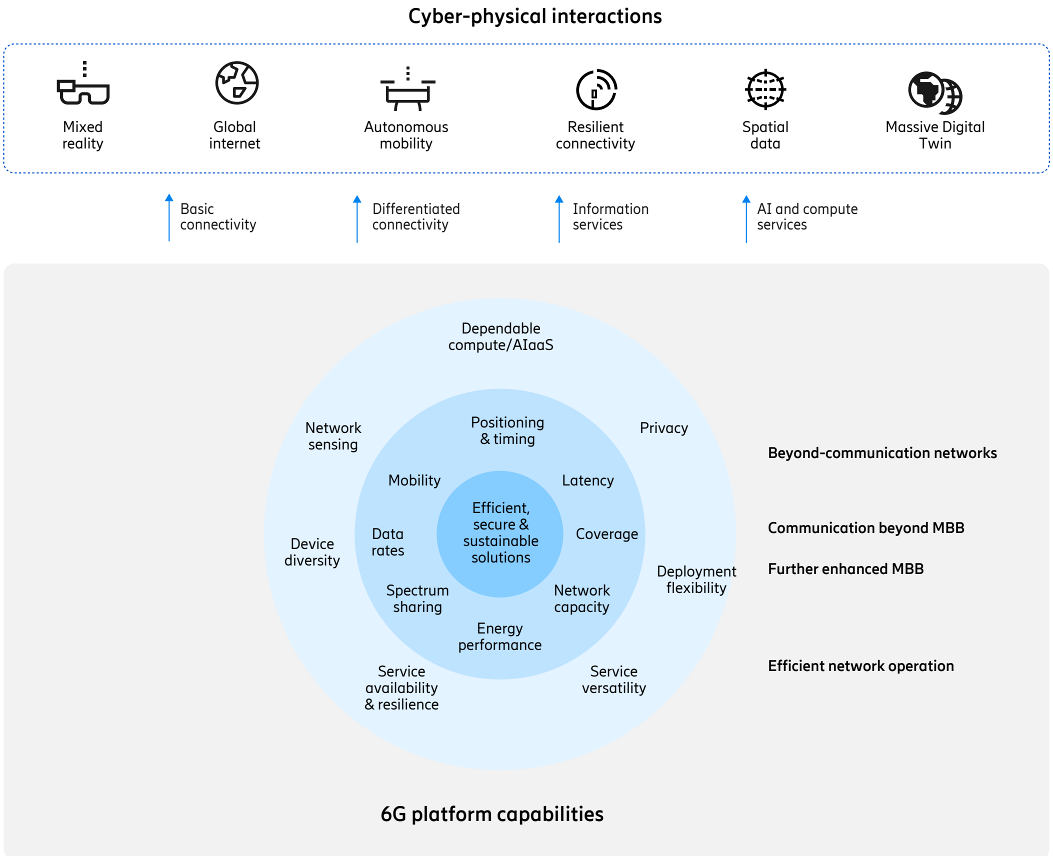


Figure 1: Examples of use cases in the cyber-physical world of 2030 served by the 6G Platform capabilities

The mentioned use cases will be enabled by improved and new capabilities in the 6G networks, as illustrated in Figure 1, importantly enhanced communication links and coverage in 6G networks, along with an expansion into beyond-communication add-on services, broadening the scope of networks into multi-purpose 6G platforms. At the same time, network operation should be improved with a focus on energy efficiency and simplifications, of importance both from cost and environmental sustainability perspectives. The rest of the paper will outline these improvements with 6G, enhancing 5G solutions and expanding further.

6G network platform

The use cases in Figure 1 entail increasing demands on performance, efficiency, sustainability, and privacy of the underlying communication networks, but also have needs for services beyond communication. As these are instrumental in enabling desired consumer experiences of applications, 6G can expand into a platform offering a wider range of services. Applications in the cyber-physical era will benefit from a richer and easier access to network capabilities through service APIs (Application Programming Interfaces), and the ability to use them in national networks or globally through aggregators.

A platform is a system that enables other businesses to share information, or build applications or content on top of it. Communication services are a key foundation for all forms of digital businesses and will continue to evolve towards tailored and programmable connectivity services leveraging customized quality of service guarantees and network slices. With the advent of 6G, networks will continue their journey towards becoming a platform that provides services going beyond connectivity, building upon the foundation laid by 5G:

Information services: The network contains and can produce information with the potential to significantly improve applications. This includes sharing network information for performance prediction and preventing mobile fraud; using positioning information to enrich existing device location sensors, including global navigation satellite system (GNSS); and providing sensing data from the radio access network (RAN) to applications in need of updated spatial data of a specific environment.

AI and compute services: the expected AI native architecture of 6G will include many functions in the network related to AI and distributed handling of large amounts of data. Some of these functions might not only be useful to operate the network but could also be exposed as services like Artificial Intelligence as a service (AIaaS) to the benefit of external parties, for example, supporting digital twins. With the same rationale, computing resources in the network could be exposed to mobile applications, for instance to offer dynamic computational device offloading to improve both device behavior (e.g., such as battery, heat) and user experience.

The 6G platform [6] as depicted in Figure 2, is a network platform built on 6G technologies and services. Based on these assets, the platform has capabilities for rich service interaction and assurance towards application service providers (ASPs), enterprises, and consumers. Communication service providers (CSPs) are in a good position to offer crucial digital services in their current local markets. However, reaching a global market has historically proved challenging. By leveraging aggregator platforms, such as network aggregators like Ericsson GNP or existing Cloud providers acting as service aggregators, CSPs can offer global services to a broad customer base and leverage developer communities. Scaling 6G-empowered applications through distribution channels that have global reach, that is, application platforms such as app stores, would create synergy effects both for telecommunications and the App economy – covering applications and application platforms, in-app purchases, third party ads, and so on – spurring further growth in both sectors.

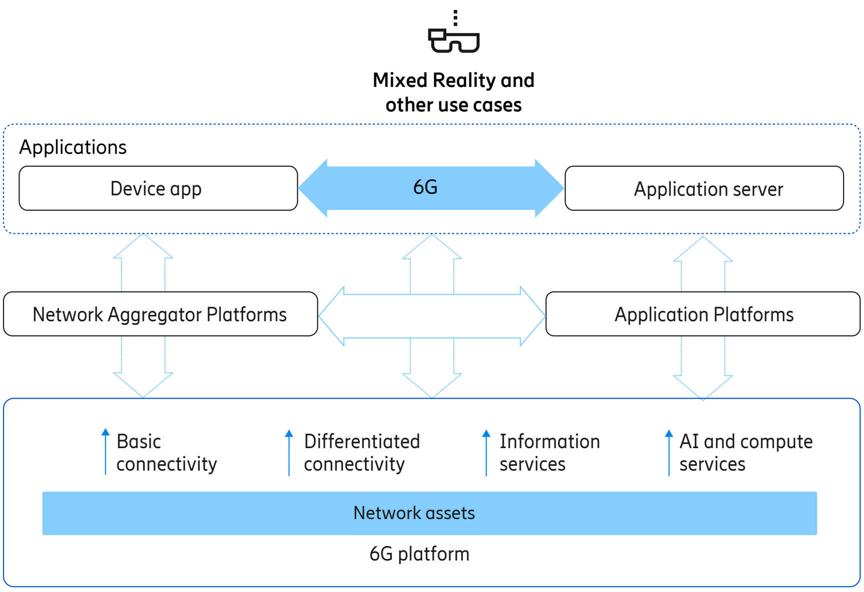


Figure 2: The 6G Platform and its role in realizing cyber-physical use cases

APIs are the means for the 6G platform to expose its capabilities to application developers. In the era of 6G, application developers will have access to a much wider range of APIs. These APIs will build on capabilities from devices and servers, multiple network domains, and services beyond just connectivity. The 5G exposure solution continues to evolve to support the 6G requirements, that is, to provide new and customized APIs more easily and rapidly. We see that service composition will be increasingly used to facilitate customized APIs. Automation support by the 6G platform will therefore be crucial to foster rapid service innovation and creation as well as to simplify life cycle management of APIs.

For services offered with performance and reliability commitments, typically expressed in a service level agreement (SLA), there is also a need for mechanisms to assure that they are fulfilled. Depending on the customer type, the network platform would provide customized SLAs, or template-based SLAs through a service-portal or from within an API. The introduction of intents provides a structured way of capturing both service and network requirements and their relative importance, including requirements from SLAs needed for service setup and assurance.

To give a concrete example, different 6G platform components could work together to support an immersive mixed reality (MR) application based on a lightweight head mounted display (HMD). To facilitate smooth application performance, the 6G platform will assist devices and ASPs with several services (see Figure 2): Tailored connectivity with user/application specific performance guarantees; positioning and environmental sensing information from the network to enhance existing device sensors for improved localization and mapping; and compute services to seamlessly offload and execute compute-intensive tasks (for example, map optimization or object detection [7]) within the device-to-cloud compute continuum, improving both HMD behavior and application quality of experience (QoE).

As illustrated in this example, such an immersive experience requires a complex interworking of many services. The user, however, will judge the quality of the overall experience mostly by the steady and smooth interaction of virtual objects with the real world as well as the convenience of carrying the HMD in a mobile scenario. Good QoE will thus reflect favorably on the ASP's offering. ASPs and consumers have an incentive to use the best communication and other services to reach that QoE.

6G architecture

The 6G architecture needs to accommodate new functionalities, for example, improved explicit network–application interactions and capabilities beyond traditional communication such as integrated sensing and AI capabilities. At the same time, it should facilitate a smooth introduction of 6G. The smooth introduction is best achieved by reusing and evolving the 5G core network (5GC) functions to support a 6G radio access network (RAN), as illustrated in Figure 3, and by adding logical functions and interfaces where necessary and business relevant. This approach makes existing 5G features such as support time-sensitive and reliable communication, available also to 6G UEs connecting through the new 6G RAN. Moreover, such smooth introduction will enable continued use, and further development of existing mechanisms for monetization, for example, network slicing and quality of service (QoS), thus helping CSPs to monetize on 6G from day one.

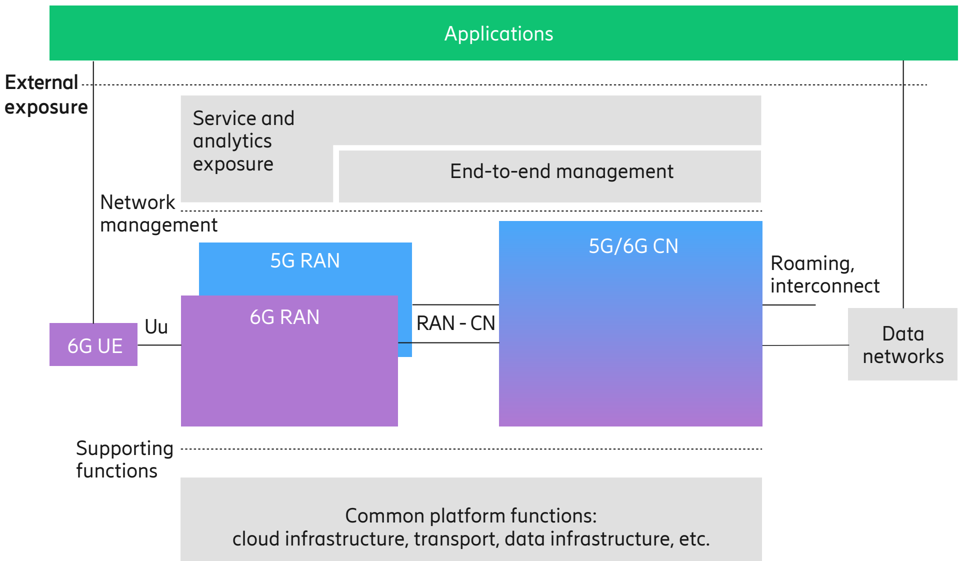


Figure 3: The 6G architecture

To simplify the migration to 6G and unify the ecosystem, the 6G architecture shall only support a standalone operation, that is, a 6G UE connects through the 6G RAN to the evolved 5GC. Every interface comes with a cost for standardization, implementation, integration, and testing. Therefore, an interface materializes in practical deployments only if the technical- and business-value of the functional split between the entities on each side of the interface justifies the interface's cost.

In the RAN, the radio interface is the most important multi-vendor interface and is certainly worth its cost. Besides the radio interface, the RAN - CN interface is a well-established multi-vendor interface. It is envisioned to have a similar functional split as today´s interface for functions evolving from 5G. Furthermore, this interface should be enhanced to suit cloud deployments where among others, reliability, availability, and resilience of the connectivity services must be ensured.

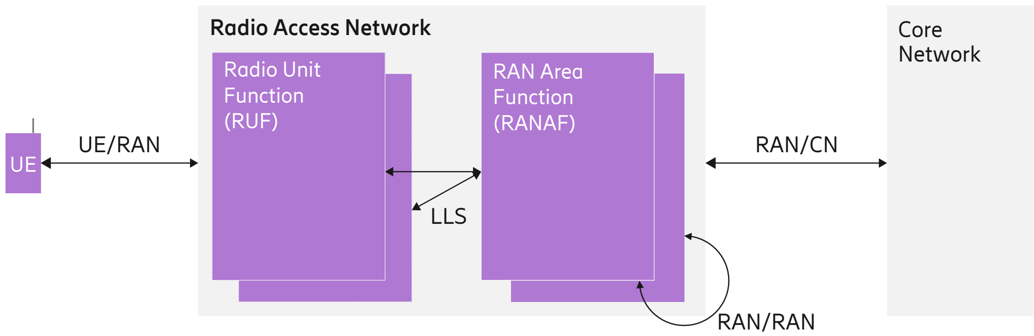


Figure 4: Interfaces within the RAN and towards the CN

In the RAN, the so-called lower layer split (LLS) interface connects the radio unit (RU) to the RAN Area Function (RANAF), see Figure 4. While the LLS is demanding in terms of throughput, latency, and jitter; it has the benefit of being stateless and loosely coupled. The RANAF maintains the knowledge about the user equipments (UEs) and their state. It instructs the radio unit (RU) dynamically when and where to transmit or receive. The RU executes the commands but does not need to maintain per-UE states. This makes it feasible to realize the LLS as a multi-vendor interface. To ensure seamless mobility between RUs connected to different RANAF, there is a need for an interface like the Xn interface in 5G.

Introducing new use cases in 6G, as described earlier, may require new functionalities in the RAN and CN architecture as well as in other domains. For sensing both the RAN and CN are expected to have a part in the functions to control sensing tasks and process the sensing data. The functions for exposing the sensing service and its result to consuming applications, on the other hand, is expected to be done as part of the CN framework for exposure. Improved explicit network application interactions may lead to new interfaces being required, exposing network capabilities to applications through the control plane, in-band in the user plane, and/or as northbound APIs. These interfaces might be with either

RAN or the CN. To be successful, such new functionalities and interfaces must be based on functional splits and interfaces that can be implemented, tested, and deployed at a cost that is justified by the business value of the targeted use case and service.

To support the increasing use of data for multiple purposes, for example, integrated, possibly distributed, AI-based solutions (as described in next chapter on AI native aspects), and observability for service assurance, there is a need for a common data plane enabling cost-efficient collection, transport, potentially storage, and governance of network data.

For more details on 6G architecture, see [8].

AI native aspects in the 6G architecture

AI will have a central role in realizing the 6G system, both as enabler to increase efficiency of existing network features, and introduce new features that were previously not possible to implement using non-AI technologies. To this end, we introduce AI native, defined as "the concept of having intrinsic trustworthy AI capabilities, where AI is a natural part of the functionality, in terms of design, deployment, operation, and maintenance" [9].

AI is pervasive throughout the entire AI native architecture (see Figure 5). This can be achieved in new products by considering AI from the start. Legacy products can evolve into AI native where this makes business sense. We see four main aspects of an AI native architecture (see also Figure 6):

1. Intelligence everywhere
2. A distributed data infrastructure
3. Autonomous operations
4. AIaaS

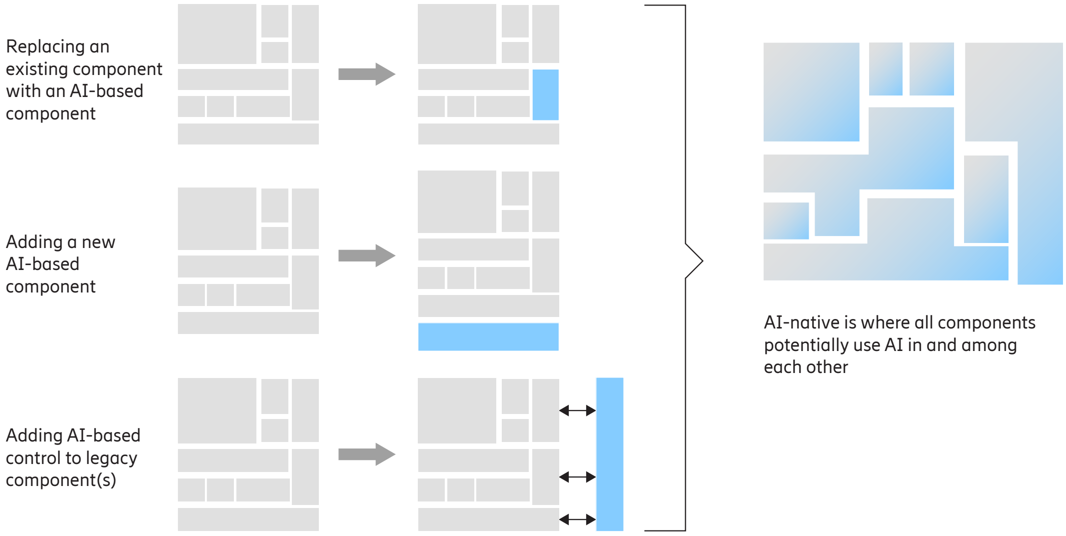


Figure 5: An AI native architecture

The aspect of intelligence everywhere is about executing AI workloads wherever it makes sense based on a cost-benefit analysis. That means, in every network domain, on every layer of a stack, on every physical site from central to edge sites, and possibly even on mobile devices. This also implies that AI execution environments need to be available everywhere, and AI training environments might be co-located if needed.

We envision a fully automated trustworthy model lifecycle management, constantly improving and following data changes, to achieve system-wide end-to-end gains. It ensures that models operate accurately and robustly even with unseen data. It decides what model version to use and where and when to perform (re)training. Models may require data spanning several layers and network domains, which may imply that layer and domain borders blur. Models with similar input features might be combined. Models might also be trained jointly with other models solving similar tasks using different inputs to benefit from multi-modality, thus reducing the number of models that need to be maintained. Additionally, models may be split into different locations, and trained in a collaborative fashion to preserve data privacy and utilize resources more efficiently by reducing the volume of data that is exchanged [10].

Intelligence everywhere is interconnected with the aspect of a distributed data infrastructure (see also the chapter on 6G Architecture). Executing and (when needed) training AI models in any location is only possible if data and necessary compute resources are available everywhere. Generative models may be used alongside network observations as data sources. Observed data may have a best-before date or legal constraints. The sheer volume of data could impose limitations, restricting when and where data can be utilized. Data streams might need to be processed or combined. Data observability needs to be flexible to adapt to the requirements of the data consumer and the available resources of the data producer and the transport infrastructure. This implies that the data infrastructure and the model orchestrators need to interact; sometimes data can be transported to the intelligence, and sometimes it is more efficient to bring the intelligence closer to the data.

There is a need to further automate the management of the network. Instead of introducing new manual operations (where humans decide what to do and how) or automated operations (where humans design workflow executions), the aim should be for fully autonomous operations [11]. Instead of instructing the system on specific actions to take, humans would express requirements to the system and supervise that those requirements are fulfilled.

Intents are a vital technology for achieving autonomous operations. We define an intent as a "formal specification of all expectations including requirements, goals, and constraints provided to a technical system" [12]. This includes requirements from diverse customers (for example, derived from SLAs) and the service provider for expressing its business policies. Various AI technologies, including Generative AI and Language Models, may be used to implement intent handling.

The AI native architecture aspects of intelligence everywhere, distributed data infrastructure, and autonomous operation require novel functions in the network related to AI and data handling. Some of these capabilities may be exposed as services to customers of the service provider; also known as AIaaS [13]. Examples include a model training or execution environment, or data exposure. Exposing such services turns the network into a platform for innovation.

For all the four main aspects of AI native, standardization has already started. This includes standardization of distributed intelligence in 3GPP SA2 (intelligence everywhere aspect), and standardization of intent interfaces in TM Forum, O-RAN, and 3GPP SA5 (autonomous operations aspect). We expect standardization of the AI aspects to continue into the 6G timeframe.

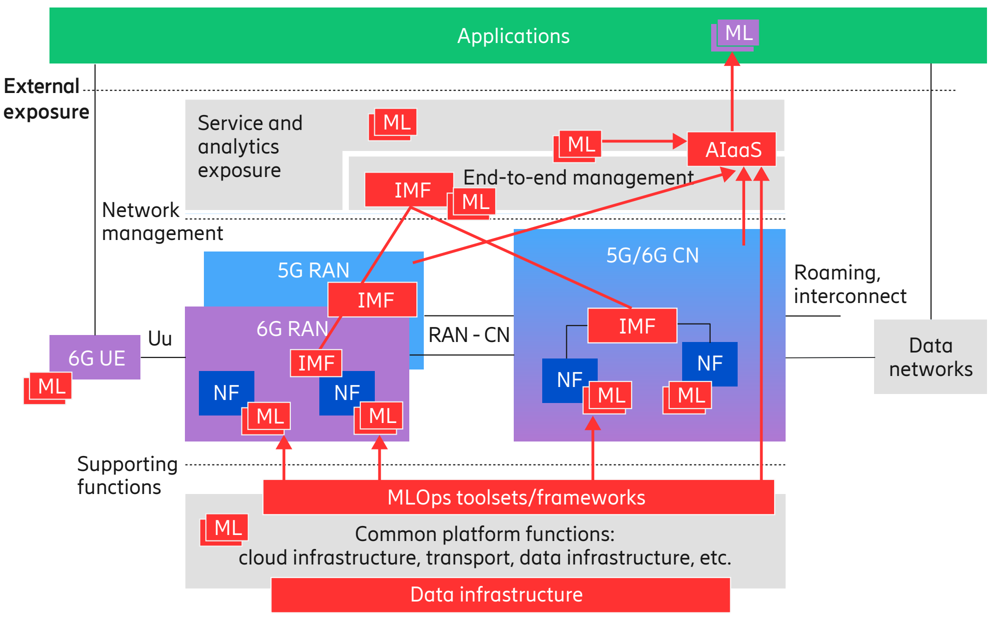


Figure 6: The four aspects of AI native shown in the functional architecture

6G radio access

Spectrum and multi-RAT spectrum sharing

Spectrum is a fundamental aspect of any radio-communication system and an important asset for a communications service provider. A wide range of frequency bands are targeted for 6G [14], including the FR1 and FR2 ranges, see Figure 7. FR1 is extremely important for wide area 6G coverage. Unfortunately, there is very little, if any, pristine spectrum in this range and 6G therefore needs to share spectrum with the previous generation. This is commonly referred to as multi-RAT spectrum sharing (MRSS). Highly efficient MRSS between 5G and 6G is essential and should be an integrated part of the 6G design from the start. Fortunately, the ultra-lean design of 5G implies that 5G-6G sharing can be made very efficient with an overhead of a few percent at most. Sharing between 6G and catM/NB-IoT is less critical and can be done through semi-static split between the RATs as the IoT technologies use a relatively small amount of bandwidth, which can be concentrated to one or a few carriers. Sharing between 4G and 6G is not seen as critical; by the time 6G is introduced 4G has largely been replaced by 5G.

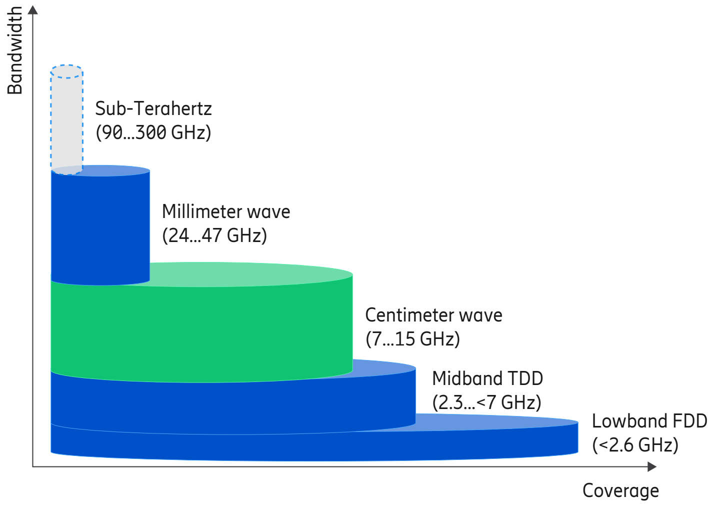


Figure 7: Spectrum for 5G

In addition to FR1 and FR2, the centimeter wave spectrum, 7 to 15 GHz, is a promising band not previously used for cellular communication. This new band can provide additional capacity for 6G and, when combined with very large antenna arrays and beamforming, downlink coverage is on par with existing midband deployments.

The sub-THz range, 100-300 GHz, on the other hand has challenging propagation conditions and the RF technology is not yet mature enough. Therefore, this frequency range is not expected to be part of the first 6G release.

Guiding principles for a high-performing 6G radio-access technology

A guiding principle for the design of the 6G radio-access technology (RAT) should involve drawing insights from the 5G experience and focusing on areas that can be improved compared to 5G. For example, the 5G OFDM waveform should be retained as no significantly better waveform has emerged. Furthermore, MRSS, an essential component of 6G, is significantly simpler if the 5G and 6G waveforms and numerologies are aligned. A similar discussion can be made around coding – low-density parity-check code (LDPC) and polar codes for data and control channels, respectively, are good choices in 6G as well.

In other areas, there is more room for improvement. Traffic patterns are highly bursty, and mechanisms to rapidly scale up the effective transmission bandwidth are needed – in other words, highly dynamic adaptation of transmission characteristics are important. Efficient handling of wide carrier bandwidths, highly dynamic and advanced multiple input and multiple output (MIMO) schemes, and a fast carrier-aggregation-like scheme for spectrum aggregation are all crucial for high-performing 6G networks. Furthermore, all aggregated carriers should be treated equally, that is, without being separated into PCell and SCells, to allow for increased robustness and a more flexible "mix and match" strategy for using the uplink and downlink spectrum, resulting in better exploitation of the service provider's spectrum assets. Together with redesigned uplink control signaling allowing for more efficient scheduling decisions, this can result in significant performance advantages.

As an example of the AI native intelligence everywhere aspect mentioned in the previous chapter on AI native aspects, Rel-19 of 5G Advanced specifies normative support for Life Cycle Management (LCM) to enable AI/ML at the air interface for a selected set of use cases, including at least positioning and beam management. This functionality likely propagates to 6G, extended to more use cases. 6G will also see application of AI/ML functionality to the shared data channel, likely in the form of power amplifier post distortion techniques, deep-learning-based receivers, and simple end-to-end- learning-based schemes.

5G already supports positioning of active devices, 6G will extend this to sensing of passive objects. On the radio access layer, this will be supported by enabling transmission of sensing reference signals. For efficient multiplexing with communication signaling, sensing reference signals should also be based on orthogonal frequency division multiplexing (OFDM). Transmitted sensing reference signals will reflect on targets and in the environment, and processing of the received reflections allow target detection, environmental sensing, etcetera.

For more details on 6G RAN design, see [15].

Energy performance

Energy efficiency is one of the most important areas for 6G, both on the network side and in the device itself. In 6G, the ultra-lean approach of 5G should extend beyond the time domain to include the frequency and spatial domains as well, see Figure 8. If, on a dynamic basis, a certain carrier or transmission point is not needed for data transmission in the network, it should be possible to put it to sleep and rapidly wake up when needed. This can be achieved by having separate signals for idle and connected mode UEs, unlike the 5G design where basically all transmissions depend on the synchronization signal block (SSB). It also enables separate optimization of idle and connected mode functionality, which have quite different requirements.

Device energy efficiency is equally important — a long battery lifetime is one of the most important aspects. Wake-up signals are one example of a feature that should be integrated in 6G from the beginning. Another example is dynamic bandwidth adaptation: A device uses a fraction of the overall carrier bandwidth to monitor scheduling assignments/grants and to receive small packets, and dynamically switches to a larger bandwidth when needed.

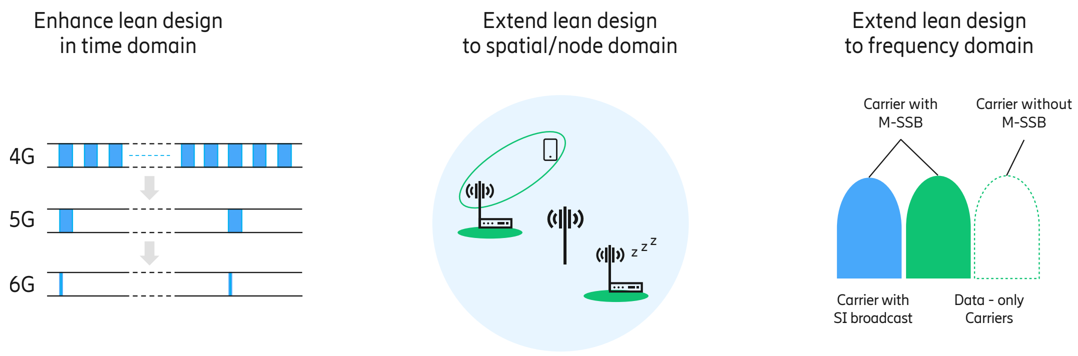


Figure 8: Ultra-lean design extended to the spatial and frequency domains

6G infrastructure and devices

Cloud infrastructure

As telco cloud infrastructures evolve towards pure cloud native-based solutions, network functions are deployed as cloud workloads (for example, in containers) to provide the necessary network agility and innovations required for 6G. For this purpose, cloud infrastructures are becoming highly heterogeneous, proposing a combination of advanced CPUs and hardware (HW) accelerators to further enhance the overall performance and energy efficiency of existing and future compute-intensive workloads. This combination of newly emerging hardware can efficiently perform specific tasks to fulfill stringent 6G requirements, allowing energy efficiency and sustainability solutions without compromising on performance. Additionally, it is fit to enable AI technologies to empower an innovative 6G ecosystem. This means that 6G standards need to be designed to accommodate the integration of the latest hardware and software innovations in future network implementations, including commodity HW technologies (such as CPUs and new instructions, special-purpose accelerators, smart networking devices, server and interconnect technologies, and so on) and industry-standard frameworks and APIs (for example, DPDK, CUDA, DOCA, OneAPI, P4, Kubernetes, etcetera) enabling the adoption of the latest and most relevant HW technologies.

Computational offloading

As cloud infrastructure used for 6G will form the base of the vastly distributed mobile networks, there is an opportunity for CSPs to take advantage of these assets to offer additional services alongside communication services. An example is a computational device offloading service [16] exposed by developer-friendly APIs. Computational offloading aims at a dynamic and flexible deployment of highly granular application tasks, triggered by situational changes. This is a use case agnostic, offering sandboxed, operator-deployed compute to any application running on a mobile device such as smartphones, MR headsets, or drones. Such a service is attractive for subscribers as it facilitates prolonged battery life, reduced device heat, and improved application experiences. Offloading can also be used by application providers to differentiate their offering, by customizing application features and performance profiles, or to address privacy concerns with guaranteed data locality compared to many current public cloud alternatives. Finally, as discussed in the Chapter on the 6G network platform, richer service offerings like computational offloading allow CSPs to leverage new revenue streams by bundling it with connectivity services or other network services like sensing data or AIaaS, while reusing existing network procedures like authentication and management functions.

Integrated sensing

Integrated sensing and communication (ISAC) is an example of a service where the same network infrastructure can be used for multiple purposes, such as sharing base stations, spectrum, data planes and pipelines, and compute between sensing and communications services [17]. This is, of course, helpful from a cost and deployment perspective. Still, some additional functions need to run in the network to build a new service such as sensing. Base stations must employ specific radio signals for sensing. Control mechanisms are needed to establish sensing in the network, and processing is needed to interpret the raw data. In addition, a rich API towards applications is needed so that they can request different sensing services needed in diverse immersive and awareness types of use cases. Privacy questions are also paramount. Sensing needs to be performed in a manner that ensures the privacy of individuals as well as the society as a whole, and sensing applications should prioritize privacy assurance.

Device diversity

The diversity of device types is expected to increase further in 6G timeframe; to serve both new and evolving use cases, see Figure 9. This device diversity, which will be active concurrently in a 6G system, makes it pertinent to understand the characteristics and requirements brought about by both novel and evolving device types. This is, in addition, to shared common requirements including improved energy efficiency, enhanced user privacy, and the adoption of relevant novel spectrum such as in the centimetric range.

Pertaining to insight use cases, we envision two 6G IoT device types. First one is the low power wide area (LPWA) devices, which desirably operate in 3-5MHz, have further reduced device complexity, similar or lower coverage enhancements, reduced signaling and UE power-saving features than NB-IoT or Cat-M1. The second one, likely in a later 6G phase, is the zero energy IoT device that obtains energy from harvesting or rely on backscattering communication and thus very low powers resulting in inherently short-range.

To realize immersive experiences, the devices such as MR devices will bring about requirements [18] on guaranteed data rates with low latency (for example, less than 20 ms) driven by motion-to-photon latency and compute location considerations. The different service flows and their attributes such as compressed, or uncompressed, along with application aspects like the exact scene or its parts [19] that are processed together will drive the communication link reliability requirements. The device form factor limitations would bring unique requirements, for example, on antenna locations and the devices with compute capability constraints would benefit from compute offloading located on network edge. The latter in turn would bring forth requirements on joint compute-communication resource optimization, communication protocol optimizations, and so on. The devices with enhanced AI capability could facilitate semantic communications, reducing bandwidth use.

Inclusivity will be ensured with not only devices that support non-terrestrial network access and fixed wireless access, but also ones that are cost-effective.

The resilient use cases require similar network and device capabilities, to ensure the appropriate level of end-to-end resilience availability and reliability. Different approaches are possible, where some of them require reliability improvements for devices used for critical use cases. When the modems are unable to fulfill reliability requirements or when economies of scale are insufficient for the necessary modem adaptations, it is possible to leverage devices with redundant modems. These redundant modems can be connected to different networks and help avoid certain common points of failure. Balancing requirements between networks and devices requires close co-operation with respective vendors to evaluate suitable approaches.

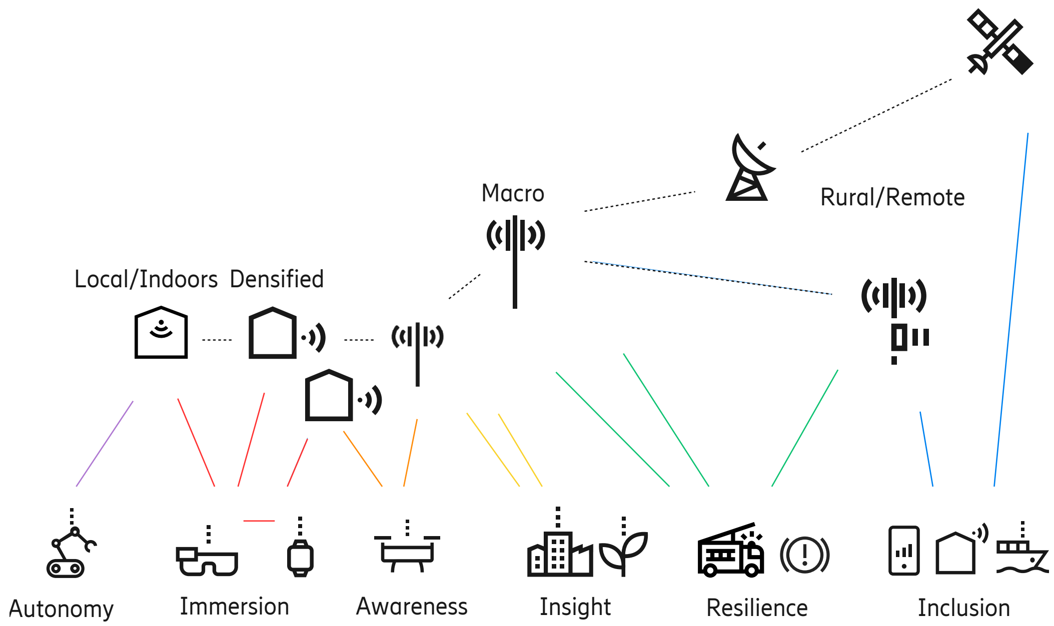


Figure 9: Device and network deployment diversity supporting many aspects of a cyber-physical world

Transport

The transport network plays a vital role in the mobile network by providing the necessary connectivity between the core network and the RAN, and between different entities within the RAN, such as RU, DU, and CU. The mix of technologies used to realize the transport network should evolve and enable CSPs to build, or upgrade existing, transport networks that are scalable, easy to maintain, smart and lean, and support the various deployments and new services provided by 6G.

The architecture of the transport network is determined by the mix of mobile network deployments, which come in many different sizes and shapes. For example, the more demanding centralized RAN (CRAN), where baseband and radio units are in different locations and are connected through fronthaul interfaces, or the simpler but likely more common distributed RAN (DRAN), where baseband and radio units are located at the same cell site and cell sites are connected through backhaul interfaces.

Furthermore, the transport networks need to concurrently support multiple services that can have strongly differentiated requirements such as eMBB, MR, or critical services. This means that the transport network also must deliver high precision network synchronization. Hence, we need transport technologies that can also handle these extreme cases when and where needed. To meet these 6G requirements, the transmission and switching technologies (packet, optical, and wireless) used in the transport network need to evolve by taking an end-to-end network perspective into account and identify solutions for network orchestration and automated management that results in a smarter network that can do more without being over-dimensioned.

Finally, it is important to understand that for most deployments, an ideal transport network is one that meets the diverse service requirements in a reliable, cost-efficient, automated, and sustainable manner, always maintaining appropriate and relevant dimensions.

RAN deployment

One of the key characteristics of 6G networks will be their ability to support a wide range of new and evolved use cases while still delivering high quality MBB and voice services, see illustration in Figure 9. This ability will depend on the availability of enhanced network capabilities that can meet the demanding requirements imposed by new services. Beyond volume, new services will require low latency and high reliability combined with high data rates. Apart from communication services, there are additional requirements related to positioning accuracy, sensing, and AI capabilities. Large-scale rollout of software features for enabling consistent low-latency and high reliability, coverage, and capacity enhancements, combined with targeted densification will be essential for meeting these requirements and providing a good user experience.

6G should be designed to enable basic versions of use cases to work on existing 5G site grids. This is important to limit initial investments. As shown in [18], many modern networks are already able to support less demanding services, like basic XR, both in terms of coverage and capacity. More demanding services will require a combination of enhanced functionality for bounded latency and increased reliability, features for enhanced uplink coverage and capacity, new 6G spectrum in the centimetric range, and densification. As network densification is complex, it will be done organically depending on the local area needs. Densification may be done by adding macro, street or indoor sites, UL-only sites, or dedicated positioning/sensing nodes. Indoor demand and performance will also influence RAN evolution for new services, especially at higher frequencies where outdoor to indoor coverage becomes very challenging. Although downlink performance in these frequencies may be compensated by higher power and beamforming, uplink performance is still fundamentally limited, and may drive indoor deployments for higher bands.

Use cases that require service guarantees in form of SLA will also require an appropriate level of deployment resilience. Depending on the requirements, resilience may need to be achieved by deployment redundancy in terms of radio redundancy (multiple carriers with non-co-located transmission points and overlapping coverage), transport redundancy, RAN and Core redundancies, for which cloud-native principles can be leveraged to improve efficiency. For some critical cases where the UE's single points of failure could have an impact, it is important to also evaluate UE redundancy.

Insight use cases, like massive twinning rely on collecting and analyzing data from many connected sensor devices in the network and using this aggregated knowledge to build a digital representation of an object, an area, or a process. A key component for many twinning applications will be the support for many simple devices/sensors that can provide the necessary data to build these digital representations of the real world.

Another key capability of 6G networks will be deployment flexibility, that is, the ability to support flexible, dynamic, and temporary deployments. In this way, the network will be able to quickly address new deployment needs related to new use cases and verticals. Non-terrestrial network solutions using satellites as a complement to terrestrial networks for coverage extension is one example that can enable improved global coverage while connectivity solutions like FWA relays and drones combined with sensors may be used in simple applications like soil humidity and plant monitoring for sustainable food production.

Security

The security in 6G evolves from 5G security, with additional security needed to support new use cases, technology, and paradigm shifts. The role of mobile networks as part of critical infrastructure comes also with new regulatory requirements on security because the availability and proper functioning of the networks are of interest to the society at large. Examples of such requirements are demands on higher security assurance and adherence to zero-trust principles [20].

Meeting 6G requirements requires combination of standardization, implementation, deployment, and operational aspects and processes. It is important to take a holistic view on the security of 6G networks to keep up with the evolving threats. These threats not only include traditional issues such as data confidentiality and fraud but also large-scale attacks on network availability, spyware attacks, and attacks on critical national infrastructure. Find more details on security in [21].

Roadmap to 2030

The focus of 6G discussions has shifted from research to pre-standardization alignment, where major regional organizations are formulating their views on the content and objectives of 6G. This includes both timeline discussions as well as the content of the first phase of 6G. There is a need to align these activities and find global consensus without making the scope too wide.

From a timeline perspective, several steps have been taken. As of 2023, 3GPP has committed to developing 6G specifications. A general expectation in wireless industry is that the first commercial deployments of 6G will happen around 2030. More of this is elaborated on in the Ericsson blog post [22], where the relation to the ITU-R work and timeline is also discussed. It is also important to note that 3GPP is not the only standard development organization (SDO) shaping the future, other SDOs and regulations will matter as well (for example O-RAN Alliance, TM Forum, and CAMARA).

6G should emerge as a global standard, offering support not only for the development of new use cases but also for addressing the widespread adoption of use cases initiated in 5G. Supporting multiple use cases is not unique to 6G, but in previous generations features targeting new verticals and deployment scenarios were added in later releases. The late introduction of these, sometimes complex, features resulted in delayed commercial availability, also considering the large development costs required to enable initially relatively small markets. Therefore, 6G should support new and evolved use cases with total cost of ownership (TCO) efficiency and sustainability in the center.

Considering this, it is crucial that the initial 6G radio access provides fundamental support for key use cases and deployment scenarios, while refraining from optimizing for the most extreme scenarios. This includes paving the way for massive immersive communication driving traffic and new requirements, low-power wide-area IoT, fixed wireless access, time-critical communication, and non-terrestrial access.

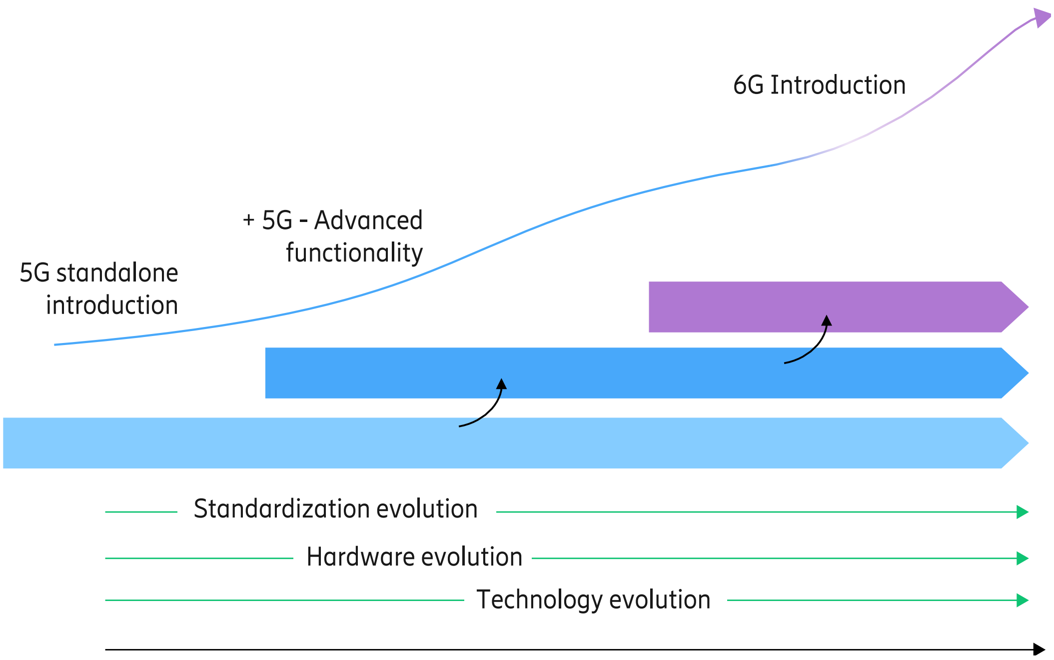


Figure 10: Network evolution to 2030 and beyond

Some steps will be crucial in paving the way for 6G, as shown in Figure 10. The first step is to move to 5G standalone supporting a rich set of MBB and Enterprise services. As a follow-up there are possibilities to add 5G-Advanced features to further expand 5G. Therefore, CSPs have the tools to monetize 5G services and capabilities including slicing, SLA assurance, observability, and network service exposure (APIs). Following these steps it is possible to quickly deploy 6G in existing bands using low overhead 5G/6G multi-RAT spectrum sharing (MRSS) enabling wide area support for new 6G features and use cases. It is therefore critical to ensure that adequate support is in place from the beginning.

Conclusion

To conclude, the following are key points for 6G:

- The day-1 baseline of 6G should offer broad support for evolved and new use cases enabling cyber-physical interactions. A diverse range of devices and network deployments should be supported.
- Programmable connectivity and beyond-communications services should be exposed as APIs to applications, creating a platform for innovation with a global market.
- An evolved 5GC should support smooth introduction of a standalone 6G RAN with a lower-layer split. The architecture should be AI native and support a high degree of automation.
- All spectrum currently used by 3GPP systems can be used by 6G, complemented by new centimeter-wave bands. Highly efficient 6G-5G sharing (MRSS) is essential in primarily FR1 and FR2.
- The 6G RAT should deliver higher performance and higher efficiency.
- Overall, focus should be put on minimizing complexity and generating value.

References

1. IMT-2030 Framework, 2023
2. 3GPP Stage-1 Workshop on IMT2030 Use Cases, 2024
3. 6G early research global takeaways, Ericsson, 2023
4. Ericsson Mobility Report, 2024
5. How generative AI is transforming telecom, Ericsson, 2023
6. The 6G platform – new opportunities for networks and applications, Ericsson Technology Review, 2024
7. XR and 5G extended reality at scale with critical communication, Ericsson Technology Review 2021
8. 6G network architecture – A proposal for early alignment, Ericsson Technology Review 2023
9. Defining AI native: A key enabler for advanced intelligent telecom networks, Ericsson WhitePaper 2023
10. Privacy-aware machine learning, Ericsson Technology Review, 2019
11. Creating autonomous networks with intent-based closed loops, Ericsson Technology Review, 2022
12. Intent in autonomous networks, TM Forum, 2022
13. AI as a service: How AI applications can benefit from the network, Ericsson 2023
14. 6G spectrum enabling the future mobile life beyond 2030, Ericsson WhitePaper 2024
15. 6G RAN – key building blocks for new 6G radio access networks, Ericsson, 2024
16. Dynamic device offloading as a 6G service, Ericsson, 2023
17. Integrated sensing and communication, Ericsson, 2024
18. Network evolution to support extended reality applications, Ericsson Technology Review 2023
19. XR in 5G, 3GPP, 2023
20. Zero Trust Architecture, NIST SP 800-207, 2020
21. 6G Security – drivers and needs, Ericsson WhitePaper, 2024
22. 6G standardization timeline and technology principles, Ericsson, 2024

Authors

Athanasios Karapantelakis is a principal researcher in artificial intelligence with a focus on neuro-symbolic systems, specifically investigating the integration of generative models with symbolic methods. He holds a Ph.D. from Royal Institute of Technology in Stockholm Sweden and has been with Ericsson since 2007.

Bengt Sahlin leads the Security Group in Ericsson Research NomadicLab in Finland. His engagement in standardization includes participating in 3GPP, ETSI, GSMA, and IETF. From 2010 to 2013 he served as the Chairman of the 3GPP security working group, TSG SA WG3.

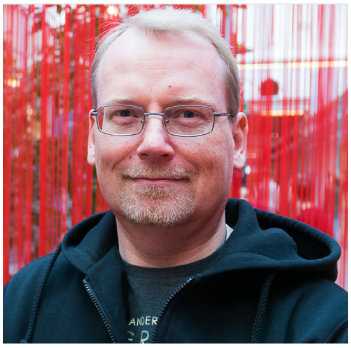
Bipin Balakrishnan is a master researcher focusing on future device technologies at Ericsson Research. He holds an M.Sc. in electrical engineering from the Royal Institute of Technology (KTH), Stockholm, Sweden.

Dinand Roeland is a principal researcher. His current research interests involve introducing AI technologies in the end-to-end network architecture, with the goal to achieve an autonomous network. He holds an M.Sc. cum laude in Computer Architectures and Intelligent Systems from the University of Groningen, the Netherlands.

Göran Rune is a senior expert in core network architecture at Research Area Networks. He holds a Lic.Eng. in solid state physics and an M.Sc. in applied physics and electrical engineering, both from the Institute of Technology at Linköping University, Sweden.

Gustav Wikström is a research leader in the area of future network technology, services and systems. He holds a Ph.D. in physics from Stockholm University, Sweden.

Henning Wiemann is a senior expert for radio network protocols. He holds a Dipl-Ing. in electrical engineering from RWTH Aachen University, Germany. He joined Ericsson Research in 2000.



Jari Arkko is senior expert in the area of Internet technology, security, and mobile networks. He holds a Lic.Tech. in computer science from Helsinki University of Technology in Finland.

Konstantinos Vandikas is a principal researcher in artificial intelligence with a focus on distributed learning applied in RAN and core network. He holds a Ph.D. in computer science from Rheinisch-Westfälische Technische Hochschule (RWTH) Aachen.

Mikael Coldrey is a principal researcher at Ericsson Research. His main research interests are in the areas of wireless communications, signal processing, advanced antenna systems, and millimeter wave communications for radio access and wireless transport systems.
He joined the company in 2006 and holds a Ph.D. in electrical engineering from Chalmers University of Technology, Sweden.

Panagiota Lioliou is a master researcher with a focus on radio features modelling and performance evaluations. She holds a Ph.D. in signal processing from Chalmers University of Technology in Gothenburg, Sweden.

Patrik Persson is holding the position as 6G program manager director at Ericsson Research. Patrik has been involved in different areas including concept development for advanced antenna systems, 3GPP RAN standardization (4G and 5G), and proprietary evolution of LTE. Patrik holds a Ph.D. in electrical engineering from the Royal Institute of Technology (KTH), Stockholm, Sweden.

Paul Schliwa-Bertling is an expert in mobile networks architecture and signaling. He holds an M.Sc. in electrical engineering from the University of Duisburg-Essen, Germany.

Per-Erik Eriksson is a master researcher focusing on research in different areas of wired and wireless transport for future technologies. He holds an M.Sc. in electrical engineering from the Royal Institute of Technology (KTH), Stockholm, Sweden.

Peter Öhlén is a principal researcher at Ericsson Research whose work centers around network and service automation, with a particular focus on how AI, intents and data can contribute to realizing the 6G vision. He joined the company in 2005 and holds a Ph.D. in photonics from KTH Royal Institute of Technology in Stockholm, Sweden.

Robert Baldemair is a principal researcher in the area of physical layer design for wireless communication systems. He holds a Ph.D. in electrical engineering from the Vienna University of Technology, Austria.

Stefan Parkvall is a senior expert based in Stockholm, Sweden. Parkvall holds a Ph.D. in electrical engineering from the Royal Institute of Technology, Sweden, and is an IEEE Fellow. He has been deeply involved in 3GPP standardization for many years and is also co-author of several popular books on 3G, 4G, and 5G such as "5G/5G-Advanced – The New Generation Wireless Access Technology".

Wolfgang John is a principal researcher in the area of compute and software, with a focus on distributed computing and edge system concepts for 5G/6G. He holds a Ph.D. in computer engineering from Chalmers University of Technology in Gothenburg, Sweden.