

Quantifying the Mutual Requirements Driving AI and 6G Co-Evolution

Brabim Mefgouda¹, Antonio de Domenico², Lina Bariah^{1,}, Clarissa Marquezan³, Riccardo Trivisonno³ and Mérouane Debbah¹*

Abstract: This paper studies the mutual requirements of artificial intelligence (AI) and sixth-generation wireless networks (6G) as they evolve together. Unlike previous wireless network generations, which added AI after deployment, 6G aims to be AI-native by embedding intelligence into its core design, interfaces, and operations. We analyze how advanced AI, such as generative AI (GenAI) and large language models (LLMs), create strict requirements for 6G networks, including low latency, high bandwidth, efficient resource usage, scalability, security, and compliance with regulations. Additionally, we identify key features that 6G networks need to support for efficient AI, such as distributed AI training, inference capabilities, dynamic resource management, programmable interfaces, and intelligent orchestration. By examining specific use cases like AI Training as a Service and LLM-based network management, the paper provides measurable insights into important performance indicators. This analysis serves as a practical guide for designing and standardizing future 6G networks optimized for AI-driven services.

Keywords: 6G networks, artificial intelligence, generative AI, large language models, network for AI, AI for network, quality of service.

1. Introduction

The co-evolution of artificial intelligence (AI) and sixth-generation cellular networks (6G) constitutes a foundational phase in the digital infrastructure, in which we are witnessing an era where connectivity and intelligence are converging at the system, architectural, and operational levels. Unlike previous wireless generations, where AI was an add-on applied post-deployment to optimize isolated functions, 6G is being designed as AI-native from the ground up, embedding AI into the core architecture, interfaces, and service models. While AI is expected to play a central role in enhancing the design, deployment, and operation of 6G networks, the opposite is equally important, in the sense that 6G

should be purposely designed to support the stringent requirements of advanced AI workloads, including among others, federated learning (FL), generative AI (GenAI) applications, etc.

A recent industry report shows that Telcos anticipate over 20% gains in revenue or cost efficiency through AI adoption, particularly in domains such as network operations, customer service, and software development [1]. AI use cases such as network capacity planning, field service assistance, and automated code generation, to name a few, are already making a notable impact in the telecom sector. However, the deployment of AI at scale introduces new pressures on the network itself, requiring 6G systems to evolve to meet new expectations of ultra-low latency, distributed intelligence, dynamic resource allocation, and energy-efficient operation [2].

This two-way relationship, AI for Network and Network for AI, requires a fundamental shift in network design. To illustrate, while AI can be utilized to optimize handovers, network behavior, resource allocation, etc., the network itself must enable programmable interfaces, reliable compute nodes, and data processing pipelines that enable distributed AI training and inference. Emerging concepts, such as AI-as-a-Service and LLM-driven orchestration, further explain how 6G will not only serve as an AI-optimized communication fabric but also as a native computational platform for serving AI workflows [3].

The aim of this paper is to characterize, in a quantitative manner, the two-way relationship between AI and 6G, AI for Network and Network for AI, by identifying the infrastructure and system demands that advanced AI workloads, such as large language models, generative AI, and distributed learning, impose on 6G networks. In parallel, we will articulate the enabling capabilities that 6G must deliver to effectively support, accelerate, and scale the deployment of AI across diverse use cases. This includes examining performance requirements such as latency, bandwidth, energy efficiency, and data accessibility, which are essential for realizing the full potential of AI-driven services in next-generation networks.

2. Measuring the Impact of AI in 6G

Although the extent of the impact of GenAI models in telecom networks in the following years is not measurable yet, several recently published studies have presented the expected and most relevant GenAI use cases [4]. A recent survey, based on data from 104 senior-level respondents from 73 communication service providers (CSPs), has identified seven families of use cases which are either being explored already today, or have short- to mid-term

¹KU 6G Research Center, Khalifa University, UAE

²Paris Research Center, Huawei Technologies, France

³Munich Research Center, Huawei Technologies, Germany

E-mail: lina.bariah@ku.ac.ae

*Corresponding Author

Manuscript received 03 June 2025, accepted 14 July 2025, and ready for publication 15 August 2025.

© 2025 River Publishers

Table 1.

Families of GenAI use cases for mobile network operators (MNOs) [5]	
Family	User Cases
Customer operations	Customer chatbot, call center agent documentation and coaching, website assistance, and predictive and personalized services.
Sales & marketing	Marketing collateral generation, personalized customer/email scripts, social media automated responses.
Network	Field service operations guided assistance, network/capacity planning, network security testing, post-mortem creation, root cause analysis.
IT/software engineering	Automated code generation and testing, automating repetitive tasks, detection of code security vulnerability.
Product innovation	Carrier billing, personalized services, voice value-added services, business-to-business, customer call services.
Internal knowledge, training & development	Evaluating new trends/developments, competitive analysis, and supply chain analysis.
Business operations	Contract, Fraud management, partner management, human resources.

potential [5]: Customer operations, sales and marketing, network, information technology (IT) and software engineering, product innovation, internal knowledge, and business operations. These families of use cases are illustrated in Table 1.

An analysis has been published presenting how GenAI models could help telcos to improve their revenues, based on a response from 130 telco operators [1]. This report identified five use case categories: customer service, sales and marketing, network, IT and support functions. Specifically, for network applications, GenAI models optimize coverage, capacity, and bandwidth, enhancing planning efficiency. They also streamline the launch of new services through efficient configuration, testing, and tuning. Finally, GenAI aids network management and orchestration by quickly resolving issues and coordinating resources, thus improving network reliability and stability [6].

2.1. Business Value of the Use Cases

Table 2 illustrates the expected impact of GenAI models per business domain of use cases [1]. This table highlights that more than 85 percent of the executives attribute to GenAI models more than 20 percent revenue or cost savings impact by business domain. Importantly, customer service, together with marketing and sales, makes up the largest share of the total impact expected in terms of business value. For instance, [1] reports that AI chatbots are anticipated to improve customer support efficiency in customer

service, potentially reducing related costs by 15 to 20 percent. Additionally, using GenAI models to summarize voice and written client interactions is expected to reduce associated costs by up to 80 percent.

In marketing and sales, MNOs use GenAI for personalized messaging, achieving over 10% customer conversion rates. GenAI also enhances network planning by structuring component data and helps IT developers double their coding productivity [7]. Support functions anticipate 30% productivity gains. Another report [8] highlights GenAI usage by telecom professionals: 57% for customer support and productivity, 48% for network management, 40% for network design, and 32% for marketing content.

2.2. AI Training as a Service in 6G Networks

In future telecommunication networks, data is generated through different sources and collected by sensors, cameras, radars and Lidars on various terminals. At the same time, with the rapid development of high-end terminals, exploiting distributed on-device processing capabilities opens the door for new AI-added-value services. In this case, 6G can provide AI training as a service in a collaborative and efficient manner by connecting data silos to enable powerful AI model production. In our vision, AI training as a service encompasses use cases for both AI for Network and Network for AI. In AI for Network, AI/machine learning (ML) models enhance network functions, while in Network for AI, networks monetize their resources and capabilities (e.g., data collection and storage) by offering AI training services to consumers through dedicated interfaces.

Figure 1 shows a high-level description of AI training as a service use case. Collected data can be transferred to the network side, where datasets are built up and used to generate powerful AI models in a centralized way. An alternative is training AI models locally instead of uploading the local raw data to the central processing centre. Making choice between the centralized or the distributed way about AI model training depends on several considerations [9, 10]. First, the decision considers the processing capabilities of the involved training entities. Second, the privacy protection of the local data should be ensured. Finally, the resources for training the AI model should be optimized.

2.3. AI Inference as a Service in 6G Networks

The rapid diffusion of AI in our society is allowing the development of numerous applications across individuals, industry, and public organizations. In our vision, 6G networks, enhanced with intelligent decision-making, will serve as a platform for delivering AI-based services. In this case, the networks can provide AI inference as a service in a collaborative and efficient manner. As for the AI training as a service, AI inference as a service is a family of use cases for both AI for Network and Network for AI.

Figure 2 illustrates three AI inference scenarios in 6G networks. In Figure 2(a), the network manages the entire AI inference workload, requiring data preprocessing to reduce latency and address privacy concerns. AI inference may be distributed across network nodes based on factors like distance and resource availability, though this introduces challenges in synchronization, security, and communication overhead. Figure 2(b) depicts AI

Table 2.

GenAI models impact on telcos by use cases [1]			
Business Domain	Share of Total Impact (%)	Share of Business Leaders by Domain (%)	Example User Cases
Customer operations	35	85	Customer-facing chatbots, call-routing performance, agent copilots, bespoke invoice creation.
Sales & marketing	35	45	Content generation, hyper-personalization, copilots for store personnel, customer sentiment analysis and synthesis
Network	15	62	Network inventory mapping, network optimization via customer sentiment analysis, self-healing via customer sentiment analysis on network problems.
IT	10	55	Copilots for software development, synthetic data generation, code migration, IT support chatbots.
Support Functions	5	10	Procurement optimization, workplace productivity, internal knowledge management, content generation HR Q&A.

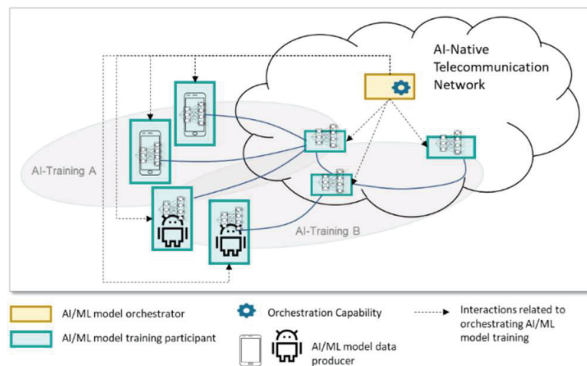


Figure 1. AI Training as a Service. AI-Training A and AI-Training B represent two instances of the AI Training as a Service.

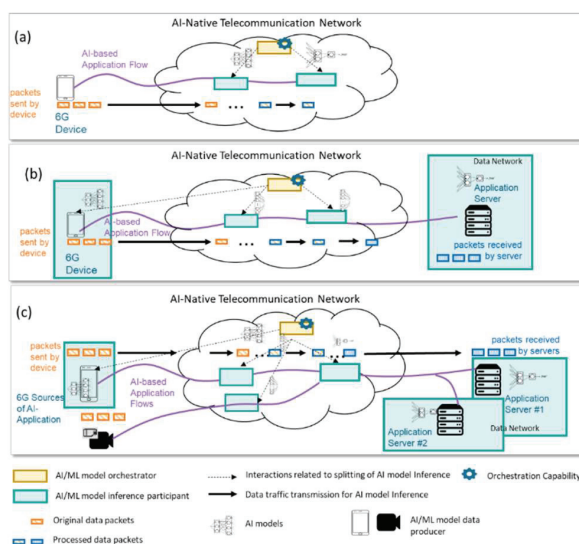


Figure 2. AI inference as a service.

inference split across the network and its environment, including users, devices, and applications. Processing part of the AI model at the user device (UE) minimizes communication needs and enhances privacy, as only processed data features are transmitted to the network. Figure 2(c) presents AI inference offloaded to the network, where an AI-enabled application accesses data from multiple sources and delivers results to multiple consumers. AI-Agents handle specialized processing, supporting applications like multi-stream fusion, while the network intelligently coordinates data sources, AI-agents, and nodes to optimize performance.

2.4. LLMs in 6G Networks

Large language models (LLMs) are advanced AI models capable of processing information and generating human-like text. In 6G networks, they offer three key capabilities that enable value-added services [11]:

- Semantic capabilities: LLMs develop an internal representation of textual data in the form of real-valued vectors called embeddings. LLMs can process and comprehend intricate information, such as the content within standard documents and infer logical conclusions from the given inputs.
- Intelligent Access to Knowledge: By understanding the specific intention conveyed through the prompt, an LLM can effectively apply its knowledge base to craft a response tailored to the user's needs.
- LLMs as Orchestrators: LLMs can utilize their knowledge to deconstruct complex tasks into manageable subtasks and deploy suitable (external) tools for each, as illustrated in Figure 3.

Integrating LLMs into 6G networks enhances AI for Network by leveraging extensive textual data, including network anomaly tickets, product manuals, and software documentation. This supports the full network lifecycle, covering planning, design, implementation, and optimization (see Figure 4). In Network for AI, 6G facilitates LLM-based AI agent coordination, model transfer, and distributed training while providing essential data collection for model development. AI agents interact with the network environment through task-specific AI capabilities for data collection, resource access via APIs, and decision dissemination.

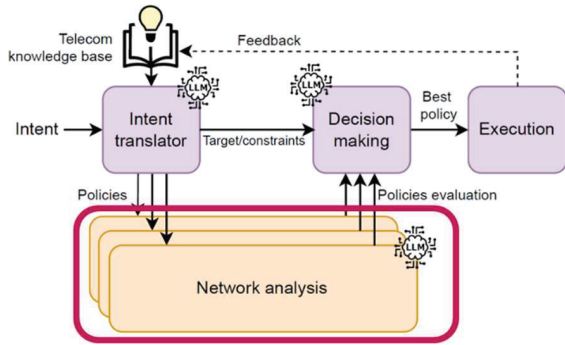


Figure 3. LLM based intent-driven network [12].

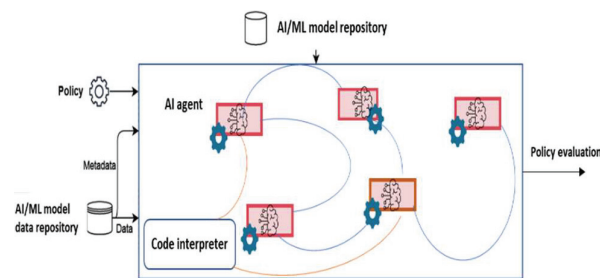


Figure 4. AI agents for network policy evaluation.

3. 6G Functional Requirements Shaped by AI

The next generation of AI, especially artificial general intelligence (AGI), is expected to play a significant role in the design, development, and operation of 6G technologies [13]. In contrast to previous generations of wireless systems, where AI was typically integrated post-deployment primarily as a tool to enhance specific functions (e.g., traffic prediction and resource allocation) [14], 6G is conceived as a fundamentally AI-native infrastructure. In this new paradigm, telecom networks will leverage AI for interacting with their environment, enabling novel value-added services, and enhancing the performance of networks and services. To realize this vision, the 6G network will host new functions and interfaces to fully support the ML model lifecycle as well as the joint optimization of connection, data collection and processing, algorithms, and models in the AI-native telecommunication network and its environment.

3.1. Data Management

The 6G network will natively support comprehensive data life-cycle management for AI training and inference, covering data collection, transmission, processing, storage, and consumption. Specifically, 6G networks will:

- Enable flexible data collection across resources, networks, and services at varying time granularities.

- Optimize data processing to enhance format compatibility, data quality, privacy, and security tailored for AI applications.
- Facilitate robust data management to support AI training, testing, and validation processes.
- Establish and maintain knowledge bases to enable retrieval Augmented Generation (RAG) capabilities [15].
- Ensure regulatory compliance for obtaining, storing, and processing private user data [4].
- Leverage user data and network-controlled information to provide innovative AI/ML-driven value-added services.

3.2. Resource & QoS Monitoring

Continuous resource and quality of service (QoS) measurements are the basis in 6G networks for successful and cost-effective AI service provisioning. Importantly, the following functions and interfaces should be introduced or enhanced:

- Functions for monitoring the status of AI resource usage.
- Function for monitoring the AI service performance requirements, such as AI model accuracy, AI service latency, or AI service density (i.e., the number of AI models operating within a given area).

3.3. 6G Interfaces for LLMs and AI Agents

To fully leverage LLMs’ semantic, orchestration, and knowledge-access capabilities, 6G networks should support interfaces exposing LLMs to:

- Knowledge bases.
- Business intents.
- Network operator policies.
- Network resource availability, user QoS, and KPIs.
- Coordination among AI agents.
- Tools such as search engines, databases, and devices.

3.4. Resource Optimization

6G networks need to jointly consider the requirements of AI value-added services together with the requirements of other (e.g., connectivity) offered services and optimize resource allocation accordingly. To achieve this goal, 6G networks integrate mechanisms and support functions for:

- Prioritizing requirements between AI Training services and other services offered by the network, considering joint communication and AI resources.
- Benchmarking available AI models with respect to a target AI service, for example, by generating and running tests to assess the service required capabilities.¹
- Determining the AI model and its configuration for a requested service offered by the 6G network. This task shall take into consideration the status of communication and AI resources.

¹ AI capabilities may include, but are not limited to, decision-making, reasoning, forecasting, clustering, classification, self-learning, acquiring contextual information, natural language processing, and complex tasks decomposition.

3.5. Resource-aware Model Training and Inference

AI consumes a notable amount of resources, including data and computers. In addition, AI/ML applications usually leverage different AI/ML models to perform the same task under different conditions with the intention to select the most accurate result [16].

6G should integrate in a native manner the capability to adapt AI models and the entire model lifecycle, including data collection and model download, to time-space varying available network resources. To attain this objective, the following mechanisms and related functions are expected:

- Intelligent mechanism for dynamically splitting the AI training and inference workload between the 6G network and its environment, based on resource availability.
- Intelligent AI training mechanism that constructs AI models able to work in different configurations to trade-off AI service performance with network resource usage [17].
- Intelligent mechanism to transfer additional layers of AI models to the inference endpoint to increase the model performance upon changes in the network conditions and AI service requirements [17].

4. 6G Performance Requirements Shaped by AI

6G systems will provide new added-value services integrating AI and leveraging telecommunication network communication and computing capabilities. Specifically, AI for Network refers to services that optimize the telecom network performance by employing AI/ML on network-specified functionalities and capabilities. This shift represents an incremental improvement in network planning strategies and a transformative advancement, providing 6G networks with real-time, autonomous decision-making, self-healing, and ubiquitous human-machine interactions. In addition, Network for AI refers to services that provide network support for AI/ML-based applications and services, such as training and inference. These newly added value services based on AI necessitate 6G networks to handle significant amounts of real-time data and adapt to changing conditions.

4.1. Ultra-Low Latency

Ultra-low latency (less than 1 millisecond) is a major requirement for AI applications that demand instant decision making, such as coordinating autonomous vehicles, performing AR surgery in real-time, and completing in-flight predictive maintenance for aviation systems [18]. To achieve this, 6G systems need specialized hardware capable of fast AI processing directly at the network edges. For instance, embedding neural processing units (NPU), which are a type of processor optimized explicitly for handling AI tasks quickly and efficiently, in 6G base stations (BSs). Consequently, AI processing occurs closer to the users and devices, significantly reducing delay. Moreover, strict QoS policies must be established specifically for these critical AI services. To enforce QoS policies, deterministic network protocols, such as IEEE 802.1 time-sensitive networking (TSN), can be employed to ensure predictable and reliable latency even in heterogeneous networks with many

connected devices. In addition, adopting new wireless waveforms and ultra-wide frequency bands will make signal delays negligible. Meeting these ultra-low latency demands ensures 6G will support current AI applications and enable future AI scenarios with even stricter real-time requirements.

4.2. Resource-Efficiency

The growing use of large-scale AI workloads is a serious energy consumption problems that arise when deployed at scale. For instance, an implementation of an advanced LLM, such as GPT-4, can consume energy equivalent to that of 100 average households in a year [19]. This high energy demand necessitates integrating energy-aware design into 6G architectures. Consequently, 6G employs neuromorphic computing methods, including spiking neural networks (SNNs), and hybrid energy harvesting, such as solar-powered nodes, to align processing with renewable energy availability and delay tolerances [20]. Moreover, Algorithmic and system-level optimizations further boost energy efficiency.

Techniques like model quantisation, structured and unstructured pruning, and network architecture search reduce parameter redundancy and computational complexity in resource-constrained wireless environments [21]. Data-centric strategies like selective sampling and prompt engineering also cut unnecessary computations while preserving model fidelity. Additionally, energy-aware network slicing enables operators to allocate resources, such as bandwidth and computing power, based on real-time carbon intensity data [22]. Collectively, these measures position 6G as a vital platform for sustainable AI deployments.

4.3. High-Bandwidth

6G networks operate in terahertz (THz) frequency bands (0.1–10 THz) with transmission rates exceeding 1 terabit per second (Tbps) [23], enabling advanced applications like tele-holograms, real-time multi-sensor fusion for autonomous vehicles, and distributed AI model development. For example, autonomous vehicles using lidar, radar, and 8K cameras generate 10–20 Tbps of raw data per hour [24], requiring efficient transmission to edge servers for immediate, safety-critical processing. Similarly, smart-city digital twin networks synchronize petabytes of urban sensor data, pushing THz-based networks to their theoretical maximum throughput. However, AI implementation at THz frequencies faces challenges due to atmospheric absorption and environmental blockage, limiting practical transmission distances to about 10 meters. To overcome these constraints, 6G incorporates AI-driven techniques such as reconfigurable beamforming, real-time channel estimation, and adaptive resource allocation. By enabling intelligent network management, AI transforms 6G bandwidth from a static resource to a dynamically optimized system, ensuring data-intensive AI applications achieve over 1 Tbps throughput while meeting stringent latency and energy efficiency requirements for next-generation wireless technologies.

4.4. Scalability

Successfully integrating AI into 6G networks requires overcoming the limitations of centralized cloud architectures, especially with

over 50 billion connected devices expected by 2030. To address these challenges, distributed AI processes data across multiple edge nodes, enhancing both scalability and privacy. Specifically, it enables parallel processing, which is essential for applications like autonomous vehicle coordination, industrial asset tracking, and augmented reality, each demanding local and distributed computing. Moreover, FL further strengthens privacy by allowing edge devices to train models locally without transmitting sensitive data, thereby improving response times and reducing network load. Additionally, hierarchical and hybrid FL models optimize bandwidth and computational efficiency by aggregating updates at intermediate nodes. Furthermore, AI-driven middleware platforms coordinate distributed AI tasks, ensuring balanced workload distribution across diverse hardware environments. Meanwhile, autonomous AI and decentralized multi-agent reinforcement learning (MARL) enhance resource allocation and enable dynamic network responsiveness. Ultimately, by prioritizing scalability and privacy, distributed AI empowers 6G networks to efficiently support AI-driven advancements.

4.5. Security and Privacy

The integration of AI into 6G networks introduces critical security and privacy challenges due to the complexities of large-scale distributed systems [25]. AI-driven edge architectures are vulnerable to adversarial, model inversion, and data poisoning attacks due to their accessibility and distributed nature [26]. AI-based security protocols evolve into advanced solutions, such as anomaly detection, automated attack response, and adaptive trust management, enhancing resilience. AI is also expanding the scope of privacy, as its implementation will necessitate the use of privacy-preserving frameworks. Collaborative AI methods will require alternative privacy-preserving techniques (e.g., differential privacy, secure multi-party computation, and homomorphic encryption) in which sensitive user data can be safeguarded when learning is decentralized. In summary, AI not only identifies important security and privacy issues, but also shapes and informs the framework under which security solutions and privacy-enhancing technologies are developed, ultimately ensuring trusted, secure, and robust 6G deployments.

4.6. Ethical and Regulatory Compliance

With the widespread use of AI for autonomous control of critical network functions in 6G, there are compelling ethical issues and regulatory compliance issues which need to be carefully addressed. AI-based decision-making will have a significant impact on essential network functions, such as resource allocation, management of access to the network, and priority of service decisions. As ethical issues of fairness, accountability, transparency, and non-discrimination come to bear open AI-based network decision-making, explainable AI (XAI) is crucial and valuable. XAI produces interpretable and transparent explanations for AI-based decisions, enabling auditing, promoting trust for users, and promoting accountability for network operators. Moreover, the pervasive application of AI within 6G networks compels regulatory bodies to develop adaptive, continuously updated compliance frameworks. These frameworks must comprehensively address

data privacy, security risks, and broader societal impacts resulting from AI integration. By explicitly embedding ethical considerations and regulatory compliance into AI systems and practices, 6G networks can ensure socially responsible and equitable technological advancement, fostering public trust and facilitating widespread adoption of future wireless technologies.

5. Conclusion

This paper presented a quantitative overview of the intertwined requirements imposed by the co-evolution of AI and 6G, which are driven by the scenarios where AI enhances the design and operation of future networks, while 6G infrastructure must support the computational and operational demands of emerging AI paradigms. Through the emergence of essential use-cases, such as AI training and inference as a service, and LLM-based orchestration, we demonstrated how 6G must evolve into an AI-native platform for supporting dynamic resource allocation, intelligent data management, and programmable services, among others. On the other hand, we outlined how AI impacts 6G performance requirements over latency, bandwidth, energy efficiency, scalability, trust, to name a few. By measuring these mutual requirements, this paper offers a foundation for guiding standardization and system/network-level design, ensuring that future networks are not just optimized by AI, but they are AI-enabling platforms at their core.

References

- [1] M. Company, "How generative AI could revitalize profitability for telcos," McKinsey, Industry Report, 2024. [Online]. Available: <https://www.mckinsey.com/industries/technology-media-and-telecommunications/our-insights/how-generative-ai-could-revitalize-profitability-for-telcos>.
- [2] Huawei, "6G: The next horizon," Huawei, White Paper, 2021. [Online]. Available: <https://www.huawei.com/en/technology-insights/industry-insights/outlook/6g/6g-next-horizon>.
- [3] L. Bariah and M. Debbah, "Telecom operators' renaissance: Seizing the opportunity in the convergence of computing and communications," McKinsey, Article, 2025. [Online]. Available: <https://www.comsoc.org/publications/ctn/pipes-intelligent-service-pipes-generational-telco-evolution-ai-could-make-reality>.
- [4] ITU-T, "TR-GenAI-Telecom Networks: Requirements and methodology for deploying and assessing Generative AI models in telecom networks," March 2025.
- [5] TM Forum, "Generative AI: Operators take their first steps," Dec. 2023. [Online]. Available: <https://inform.tmforum.org/research-and-analysis/reports/generative-ai-operators-take-their-first-steps>.
- [6] ITU-T SG13, "Recommendation ITU-T Y.3401: Coordination of networking and computing in IMT-2020 networks and beyond – Capability framework," Sept. 2024.
- [7] McKinsey, "Unleashing developer productivity with generative AI," June 2023. [Online]. Available: <https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/unleashing-developer-productivity-with-generative-ai>.
- [8] NVIDIA, "State of AI in Telecommunications: 2024 Trends," 2024. [Online]. Available: <https://resources.nvidia.com/en-us-ai-in-telco/state-of-ai-in-telco-2024-report?ncid=no-ncid>.
- [9] E. T. Martínez Beltrán, M. Q. Pérez, P. M. S. Sánchez, S. L. Bernal, Bovet, M. G. Pérez, G. M. Pérez, and A. H. Celdrán, "Decentralized federated learning: Fundamentals, state of the art, frameworks,

- trends, and challenges,” *IEEE Communications Surveys Tutorials*, vol. 25, no. 4, pp. 2983–3013, 2023.
- [10] 3GPP Technical Specification Group Services and System Aspects Management and Orchestration, “Artificial Intelligence/Machine Learning (AI/ML) management,” Release 19, March 2025.
- [11] A. Maatouk, N. Piovesan, F. Ayed, A. De Domenico, and M. Debbah, “Large language models for telecom: Forthcoming impact on the industry,” *IEEE Communications Magazine*, vol. 63, no. 1, pp. 62–68, 2025.
- [12] F. Ayed, A. Maatouk, N. Piovesan, A. De Domenico, M. Debbah, and Z.-Q. Luo, “Hermes: A large language model framework on the journey to autonomous networks,” 2024. [Online]. Available: <https://arxiv.org/abs/2411.06490>.
- [13] L. Bariah and M. Debbah, “AI embodiment through 6G: Shaping the future of AGI,” *IEEE Wireless Communications*, 2024.
- [14] C.-X. Wang, M. Di Renzo, S. Stanczak, S. Wang, and E. G. Larsson, “Artificial intelligence enabled wireless networking for 5G and beyond: Recent advances and future challenges,” *IEEE Wireless Communications*, vol. 27, no. 1, pp. 16–23, 2020.
- [15] A.-L. Bornea, F. Ayed, A. De Domenico, N. Piovesan, and A. Maatouk, “Telco-RAG: Navigating the challenges of retrieval-augmented language models for telecommunications,” 2024. [Online]. Available: <https://arxiv.org/abs/2404.15939>.
- [16] B. Taylor, V. S. Marco, W. Wolff, Y. Elkhatib, and Z. Wang, “Adaptive deep learning model selection on embedded systems,” *SIGPLAN Notices*, vol. 53, no. 6, pp. 31–43, Jun. 2018. [Online]. Available: <https://doi.org/10.1145/3299710.3211336>.
- [17] F. Ayed, A. De Domenico, A. Garcia-Rodriguez, and D. López-Pérez, “Accordion: A communication-aware machine learning framework for next generation networks,” *IEEE Communications Magazine*, vol. 61, no. 6, pp. 104–110, 2023.
- [18] B. Hassan, S. Baig, and M. Asif, “Key technologies for ultra-reliable and low-latency communication in 6G,” *IEEE Communications Standards Magazine*, vol. 5, no. 2, pp. 106–113, 2021.
- [19] J. Achiam, S. Adler, S. Agarwal, L. Ahmad, I. Akkaya, F. L. Aleman, D. Almeida, J. Altenschmidt, S. Altman, S. Anadkat, “GPT-4 technical report,” *arXiv preprint arXiv:2303.08774*, 2023.
- [20] A. Jabbari, H. Khan, S. Duraibi, I. Budhiraja, S. Gupta, and M. Omar, “Energy maximization for wireless powered communication enabled IoT devices with NOMA underlying solar powered UAV using federated reinforcement learning for 6G networks,” *IEEE Transactions on Consumer Electronics*, vol. 70, no. 1, pp. 3926–3939, 2024.
- [21] D. Sharma, V. Tilwari, and S. Pack, “An overview for designing 6G networks: Technologies, spectrum management, enhanced air interface, and AI/ML optimization,” *IEEE Internet of Things Journal*, vol. 12, no. 6, pp. 6133–6157, 2025.
- [22] H. Chergui, L. Blanco, L. A. Garrido, K. Ramantas, S. Kukliński, A. Ksentini, and C. Verikoukis, “Zero-touch AI-driven distributed management for energy-efficient 6G massive network slicing,” *IEEE Network*, vol. 35, no. 6, pp. 43–49, 2022.
- [23] A. Shafie, N. Yang, C. Han, J. M. Jornet, M. Juntti, and T. Kürner, “Terahertz communications for 6G and beyond wireless networks: Challenges, key advancements, and opportunities,” *IEEE Network*, vol. 37, no. 3, pp. 162–169, 2023.
- [24] D. Katare, D. Perino, J. Nurmi, M. Warnier, M. Janssen, and A. Y. Ding, “A survey on approximate edge AI for energy efficient autonomous driving services,” *IEEE Communications Surveys & Tutorials*, vol. 25, no. 4, pp. 2714–2754, 2023.
- [25] V.-T. Hoang, Y. A. Ergu, V.-L. Nguyen, and R.-G. Chang, “Security risks and countermeasures of adversarial attacks on AI-driven applications in 6G networks: A survey,” *Journal of Network and Computer Applications*, p. 104031, 2024.
- [26] B. D. Son, N. T. Hoa, T. V. Chien, W. Khalid, M. A. Ferrag, W. Choi, and M. Debbah, “Adversarial attacks and defenses in 6G network-assisted IoT systems,” *IEEE Internet of Things Journal*, vol. 11, no. 11, pp. 19168–19187, 2024.

Biographies



Brahim Mefgouda (Member, IEEE) is a Postdoctoral Fellow at the 6G Research Center at Khalifa University. He obtained his B.Sc. (Hons) in Computer Science in 2015 and his M.Sc (First Class) in Artificial Intelligence in 2017, both from the Department of Computer Science at the University of Biskra, Algeria. In March 2023, he completed his Ph.D. in Computer Science at the National School of Computer Science (ENSI), Manouba University, Tunisia. Dr. Mefgouda has also been a research visitor at the James Watt School of Engineering, University of Glasgow, Glasgow, UK. His research interests encompass using machine learning, generative AI, and optimizations to enhance the next generation of wireless communications.



Antonio de Domenico (Member, IEEE) received the M.Sc. degree in telecommunication engineering from the University of Rome La Sapienza, Rome, Italy, in 2008, and the Ph.D. degree in telecommunication engineering from the University of Grenoble, Grenoble, France, in 2012. From 2012 to 2019, he was a Research Engineer with CEA LETI MINATEC, Grenoble. In 2018, he was a Visiting Researcher with the University of Toronto, Toronto, ON, Canada. Since 2020, he has been a Senior Researcher with Huawei Technologies France, Paris, France. Since 2023, he has been co-leading the network energy efficiency activities within the Green Future Network Project of the NGMN Alliance. He is the main inventor or a co-inventor of more than 25 patents and authored or coauthored more than 80 publications. His research interests include heterogeneous wireless networks, machine learning, and green communications.



Lina Bariah (Senior Member, IEEE) received the Ph.D. degree in communications engineering from Khalifa University, Abu Dhabi, UAE, in 2018. She was a Visiting Researcher with the Department of Systems and Computer Engineering at Carleton University, Ottawa, Canada, and an Affiliate Research Fellow at the James Watt School of Engineering, University of Glasgow, UK. She also served as a Senior Researcher at the Technology Innovation Institute and Lead AI Scientist at Open Innovation AI. She is currently an Adjunct Professor and Scientist at Khalifa University, and an Adjunct Research Professor at Western University, Canada. She serves as the Industry Chair for the IEEE GenAINet ETI, a member of the Technical Advisory Committee – AI for Good (Innovate for Impact), a Steering Board Member of the Wireless World Research Forum (WWRF), and a Distinguished Lecturer of the IEEE Vehicular Technology Society (VTS). She was recently named among the “100 Brilliant and Inspiring Women in 6G” by the Women in 6G organization. She is a frequent speaker at international forums and industry events. She has authored/co-authored over 80 research papers and book chapters in top-tier journals and flagship conferences and currently serves as an Editor for the IEEE Transactions on Wireless Communications. Her research focuses on Generative AI for Telecom, Large Language Models, and AI for Communications.



Clarissa Marquezan currently holds the position of Principle Engineer at Network Architecture group for the Advanced Wireless Technologies Lab (AWTL) at Huawei Technologies, Munich Research Center. She works on developing research and solutions for standardization bodies in the areas of Network Architecture and Core Network. She joined Huawei Technologies in 2014. She holds a PhD and MSc in Computer Science from Federal University of Rio Grande do Sul, Brazil (achieved in 2010 and 2006, respectively).



Riccardo Trivisonno (Senior Member, IEEE) received the M.Sc. and Ph.D. degrees in telecommunications engineering from the University of Bologna, in 2000 and 2005, respectively. He joined Huawei Technologies, in 2011. He is currently the Head of Network Architecture – Research and Standardization – for the Advanced Wireless Technologies Laboratory, Munich Research Center. Over the past ten years, he has been a leading contributor to the definition and the standardization of 5G network architecture and technologies, delivering to 3GPP Releases 15-18 in the areas of architecture modularization, network slicing, network analytics, and QoS, filing more than 100 standard essential patent applications. His research focus shifted to 6G enabling technologies, since 2020. He has been the Chair of 6G-IA prestandardization WG, since 2020, a Board Member of one6G Association since its foundation, in 2021, and he was elected as the one6G Board ViceChair, in 2023.



Mérouane Debbah (Fellow, IEEE) is a researcher, educator and technology entrepreneur. Over his career, he has founded several public and industrial research centers and start-ups, and is now Professor at Khalifa University of Science and Technology in Abu Dhabi and founding Director of the KU 6G research center. He is a frequent keynote speaker at international events in the field of telecommunication and AI. His research has been lying at the interface of fundamental mathematics, algorithms, statistics, information and communication sciences with a special focus on random matrix theory and learning algorithms. In the Communication field, he has been at the heart of the development of small cells (4G), Massive MIMO (5G) and Large Intelligent Surfaces (6G) technologies. In the AI field, he is known for his work on Large Language Models, distributed AI systems for networks and semantic communications. He received multiple prestigious distinctions, prizes and best paper awards (more than 35 best paper awards) for his contributions to both fields and according to research.com is ranked as the best scientist in France in the field of Electronics and Electrical Engineering. He is an IEEE Fellow, a WWRF Fellow, a Eurasis Fellow, an AAlA Fellow, an Institut Louis Bachelier Fellow and a Membre émérite SEE.