AI/ML FOR RADIO ACCESS NETWORKS: ENABLING SELF-EVOLVING AND INTELLIGENT CONNECTIVITY

Dr. Evren Tuna

ULAK Communications Inc.

Prof. Dr. Hüseyin Arslan

Istanbul Medipol University

Attribution: Tuna, E., & Arslan, H. (2026). AI/ML for radio access networks: Enabling self-evolving and intelligent connectivity. In H. Arslan, A. Taşdelen, E. Kuzucu Hıdır, & O. A. Çetin (Eds.), *Artificial intelligence and national technology reflections* (pp. 151–169). Turkish Academy of Sciences. https://doi.org/10.53478/TUBA.978-625-6110-86-1.ch07

10.53478/TUBA.978-625-6110-86-1.ch07

AI/ML FOR RADIO ACCESS NETWORKS: ENABLING SELF-EVOLVING AND INTELLIGENT CONNECTIVITY

Dr. Evren Tuna

ULAK Communications Inc.

Prof. Dr. Hüseyin Arslan

Istanbul Medipol University

Abstract

The transformation of radio access networks (RANs) in the age of artificial intelligence (AI) reflects a broad shift shaped by domain knowledge from both wireless communications and computer science. As mobile networks evolve toward the sixth generation of mobile systems (6G), AI is becoming a native, in-network capability, embedded directly into the architecture, managing its own model lifecycle, and operating with full awareness of computational load, processing constraints, and energy budgets. This chapter examines how growing service diversity, rising traffic volumes, and increasing network complexity have exposed the limitations of traditional design and optimization approaches, creating a need for predictive, adaptive, and self-evolving network behavior. First, the evolution of mobile networks is reviewed to explain why 6G must be designed with AI as an intrinsic component and how AI/machine learning (ML) transitions from an external aid to a fully native capability in next-generation systems. Building on this foundation, the chapter highlights three representative research areas-prediction, optimization, and network management-by examining selected studies that are frequently emphasized in academia and industry. Next, the chapter discusses how standardization bodies address AI/ML capabilities across specific use cases through common architectures, unified data structures, lifecycle management frameworks, and well-defined procedures for training, inference, and deployment. This structured perspective links early-stage research to standardized, deployable solutions. Finally, the chapter summarizes the scientific and operational challenges on the path to AI-centric 6G, including generalization in non-stationary environments, privacy-preserving learning, multi-vendor interoperability, and ultra-low-latency execution. By integrating insights from research, architecture, and standardization, the chapter contributes to the book's broader vision of self-evolving, intelligence-driven connectivity.

Keywords

Wireless Communication, Radio Access Network, AI/ML, 3GPP, O-RAN

Introduction: Towards AI-Native Intelligent Connectivity

AI has become a foundational driver of technological transformation across a wide range of domains-including computer vision, robotics, and wireless networks. The evolution of mobile networks is shaped by the emergence of new types of services, the diversification of service demands over time, the integration of different vertical technologies, and the architectural refinements required to manage the increasing complexity of future systems.

As we move toward 6G, research and standardization efforts are converging across multiple disciplines, covering architectural developments, physical-layer procedures, and end-to-end network operations with a growing emphasis on predictive and autonomous capabilities. At the same time, advancements in computing and semiconductor technologies have accelerated the transition of network management from traditionally reactive approaches toward proactive and increasingly autonomous operation (Balmer et al., 2020). In parallel, the enablers of AI technology-increased data availability, improvements in computational capability, and the development of new algorithms-have positioned AI as a cornerstone in the evolution of mobile networks.

Rather than serving isolated functions within specific components, AI/ML in modern mobile networks is viewed as a native capability that spans all layers of the system-from physical-layer signal processing to high-level network management. This broad integration brings a set of scientific, architectural, and operational questions. First, AI/ML-driven solutions must be evaluated against both non-AI/ML and alternative AI/ML methods in terms of performance, computational load, energy efficiency, and implementation complexity. Second, the placement of AI/ML tasks, whether at the user equipment (UE), within the network, or in a collaborative distributed form, requires careful analysis. Third, developing AI-native, appropriate data flows and interfaces within the architecture is essential to communicate the measurements and information required for the AI/ML solution. Finally, integrating AI/ML requires strong guarantees for privacy, security, and model trustworthiness, as these functions may introduce new risks related to data exposure, unauthorized access, or model-level vulnerabilities. As these capabilities mature, AI/ML is evolving mobile infrastructures into intelligent, self-managing, adaptable, and resilient systems that can support the transformation needs of diverse vertical industries (Rashid & Kausik, 2024).

An extensive research landscape is actively investigating diverse AI-driven use cases across physical, Medium Access Control (MAC), and network layers, highlighting the richness of current scientific efforts. Complementing this, international standardization bodies and industry alliances are shaping unified frameworks that emphasize reliability, generalizability, and lifecycle management. In parallel, the research community is expanding its focus to encompass a wide range of emerging use cases, reflecting the increasing scale and maturity of current investigations. At the same time, major standardization groups, including the Third Generation Partnership Project (3GPP) and the O-RAN Alliance, are developing common frameworks for future intelligent networks, with a particular focus on reliability, generalizability, and end-to-end lifecycle management.

This chapter presents the evolution of the RAN across mobile generations in the emerging age of AI. Section 2 presents the evolution from Long Term Evolution (LTE) to 6G and explains why AI/ML has become a necessity for mobile networks. Section 3 examines how AI/ML has progressed from research concepts to structured elements in standardization. Section 4 discusses the scientific and deployment-related challenges associated with the use of AI/ML in wireless environments. Together, these sections provide a structured view of the technical foundations, requirements, and open questions that will shape the development of the AI-native mobile networks.

1. Evolution of Mobile Networks: Why 6G Must Be AI-Native

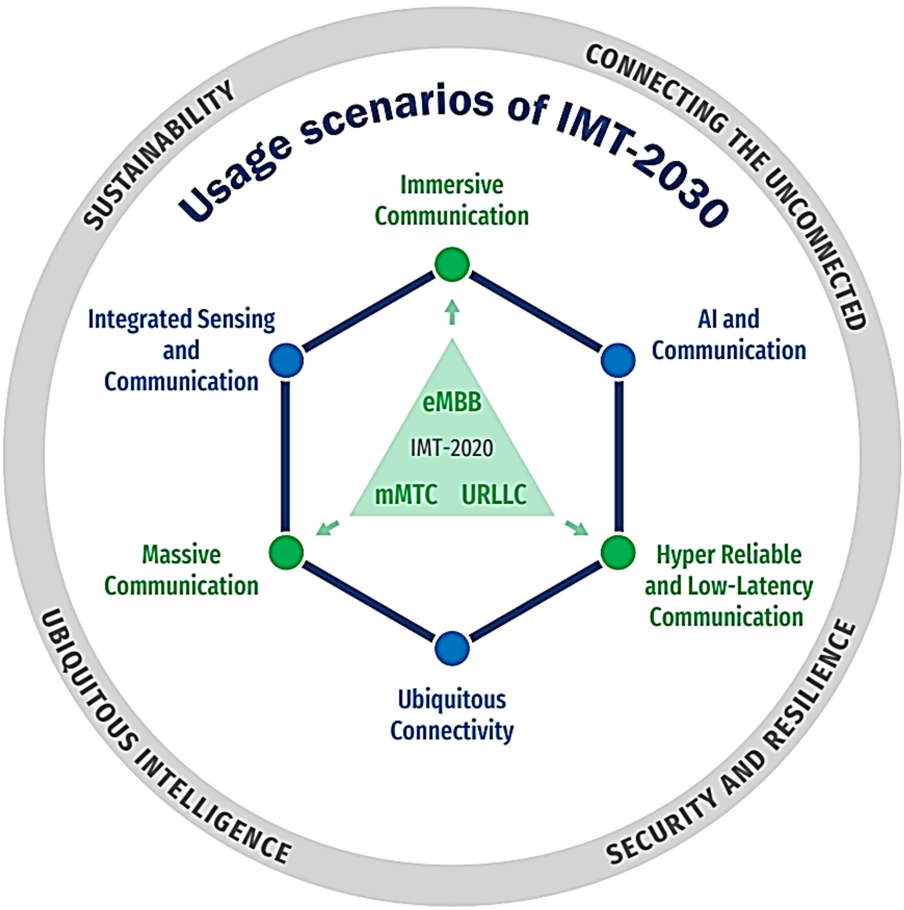
To understand why AI has become an essential component of today's wireless networks, it is helpful to trace the evolution of mobile networks, from LTE through fifth generation mobile systems (5G) and toward the emerging vision of 6G. Defined by the International Telecommunication Union (ITU) under the IMT-Advanced framework, LTE adopted a fully IP-based architecture, along with a flexible and high-efficiency radio interface, enabling globally scalable, data-centric mobile broadband (ITU-R, 2008). Although the use of AI/ML was still limited during this generation, the concept of Self-Organizing Networks introduced early forms of automation, such as predictive fault detection, load balancing, and automatic parameter configuration, laying the groundwork for adaptive optimization in mobile networks (Zhang et al., 2019; Kaur et al., 2021).

Defined by the ITU under the IMT-2020 framework, 5G redefined mobile network services by introducing three capability pillars: enhanced mobile broadband (eMBB), ultra-reliable low-latency communication (URLLC), and massive machine-type communication (mMTC) (ITU-R, 2015). The heterogeneity of these service categories and the corresponding rise in architectural and operational complexity have made the integration of AI/ML technologies increasingly essential. Building on this need, the 5G architecture specified by the 3GPP incorporates automation and data-driven management mechanisms. 5G provides an infrastructure that enables the network to improve itself, adjust itself, and recover from problems through the analysis of data in the network and decision-making mechanisms based on these analyses.

ITU's IMT-2030 vision, 6G, presents a deeper integration of AI, positioning intelligence as an intrinsic element of how networks are conceived, operated, and evolved. Usage scenarios of IMT-2030 are envisaged to expand on those of IMT-2020 into broader use requiring evolved and new capabilities. In addition to expanded IMT-2020 usage scenarios, IMT-2030 is envisaged to enable new usage scenarios arising from capabilities, such as AI and sensing, which previous generations of IMT were not designed to support. Figure 1 illustrates the usage scenarios for IMT-2030. AI and Communication use cases would support distributed computing and AI applications. This usage scenario would require support for high area traffic capacity and user-experienced data rates, as well as low latency and high reliability, depending on the specific use case. Besides communication aspects, this usage scenario is expected to include a set of new capabilities related to the integration of AI and compute functionalities into IMT-2030, including data acquisition, preparation, and processing from different sources, distributed AI model training, model sharing, and distributed inference across IMT systems, and computing resource orchestration and chaining (ITU-R, 2023).

Figure 1

Usage Scenarios of IMT-2030 (ITU-R, 2023)



As the evolution of mobile networks shows, network operations-including performance monitoring, configuration management, and control functions-initially relied on manual, operator-driven processes and later shifted toward rule-based and scripted mechanisms. Over time, these reactive approaches have struggled to scale with the growing size and complexity of modern networks. The expansion of service types, the heterogeneity of RAN and core network (CN) architectures, differences in UE capabilities, and the resulting computational and energy burdens have increased the need for more proactive operational models. As a result, wireless networks have reached a stage where conventional deterministic models are no longer sufficient on their own, creating a strong demand for learning-based and adaptive intelligence.

AI/ML, therefore, advances this evolution by introducing intelligence and adaptability into end-to-end network operations (Chataut et al., 2024; Mahmood et al., 2022). AI/ML approaches, applied directly or indirectly, enable the network to manage the balance between latency, energy consumption and quality of service (QoS); for example, in this context, many decisions such as resource allocation, spectrum usage, load distribution, power settings and mobility operations can be handled together at different layers of the network and in different usage scenarios (Dulaj et al., 2025; Zhao et al., 2024). Complementing these capabilities, deep learning (DL) enables the handling of complex tasks in rapidly changing wireless environments with greater accuracy. In contrast, reinforcement learning (RL) enables networks to learn from experience and gradually improve their decision-making over time (Zhang et al., 2019; Kaur et al., 2021). Together, these advances shift the network from isolated optimizations toward a cohesive, intelligence-driven operation model.

To provide insight into the kinds of problems AI/ML seeks to solve within mobile networks, three representative research-driven application areas are frequently emphasized in the literature. First, dynamic spectrum management can contribute to more efficient spectrum utilization, maintain connection availability, and enhance the user experience using AI/ML. In addition, a recent study presents a system that estimates traffic density and interference patterns, and adjusts channel, power, and modulation settings in real-time (Sanjalawe et al., 2025). A second major line of work focuses on mobility management. Studies in this area demonstrate that handover decisions can be supported by reduced measurement load and latency, which is made possible by incorporating parameters such as signal strength and disconnection information into AI/ML models. Finally, further research demonstrates that AI/ML plays a critical role in enabling energy-efficient base station (BS) design and operation (Usman et al., 2025). The proposed system dynamically reconfigures BS and antenna-array activity based on real-time traffic conditions, minimizing energy consumption without compromising service quality.

As mobile networks continue to evolve, the underlying assumptions about how AI/ML solutions are executed and integrated within the system are also undergoing significant changes. In the past, a common approach was for such solutions to run in a network-agnostic environment, often on third-party systems, and to be integrated using shared data across existing interfaces. These solutions operated without requiring any changes to the RAN or CN architecture. However, as we move towards 6G, it becomes increasingly important to deploy AI/ML-based algorithms directly within the network, at appropriate transmission reception points, facilitate data exchange, and create architectures that enable the necessary information to pass through the relevant interfaces.

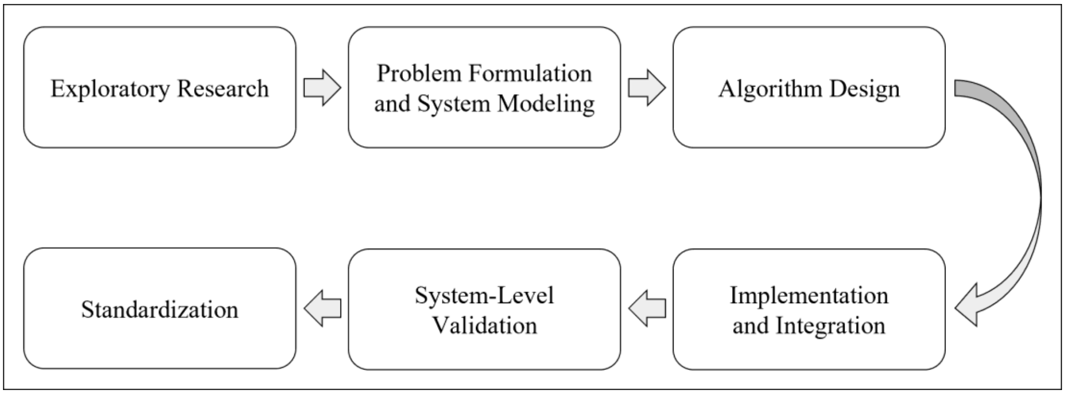
For AI/ML models to operate reliably in the dynamic wireless environment, generalization, lifecycle management, updating, and transferring challenges are among the fundamental topics of AI/ML research in 6G. This is because performance in a specific and controlled environment can differ significantly from real-world wireless conditions. Therefore, both within 3GPP and the O-RAN Alliance, intensive work is underway to standardize processes such as model monitoring, feedback mechanisms, retraining, versioning, and so on (3GPP, 2025a). These developments are paving the way for AI/ML to move beyond an idea tested solely in the research phase and become an established part of industrial network architecture (O-RAN, 2025). With 6G, AI/ML is no longer a side support but rather a fundamental element of network design itself.

2. AI as the Engine of Network Evolution: From Research Concepts to Standardized Capabilities

In the 6G era, AI/ML-driven innovations typically progress through a structured Technology Readiness Level (TRL) pipeline, beginning with exploratory research, followed by mathematical modeling and algorithm design, intellectual property development, and finally system-level validation and standardization-as illustrated in Figure 2. In the research landscape, a wide range of problems across different layers of mobile networks such as channel estimation, beam management, Channel State Information (CSI) feedback enhancement, positioning, Radio Resource Management (RRM) measurement prediction, network slicing, and so on, are being explored through AI/ML-based methods. For any of these methods to progress from a research idea to an actual deployable solution, several structured stages must typically be followed. This section first highlights several prominent studies within the research landscape and subsequently examines standardization efforts-primarily those of 3GPP and the O-RAN Alliance.

Figure 2

TRL Pipeline for AI/ML Development and Standardization



2.1. AI/ML Research Landscape for Prediction, Optimization, and Network Management

In the research landscape, AI/ML has been applied to many different problems across wireless network architecture. These studies vary widely in scope, but a large part of them can be grouped into three main areas: prediction, optimization, and management. This grouping does not cover every possible use case, but it helps organize the research directions that are most relevant for practical deployment and future standardization efforts.

Recent work on prediction covers several topics, such as traffic-load prediction, link-quality prediction, and methods that provide early signs of potential faults. These studies use measurements collected from the network to infer short-term changes in load, interference, or radio conditions. The resulting predictions are not viewed as standalone outcomes; instead, they serve as inputs that can support decisions in scheduling, resource allocation, or maintenance planning.

The network traffic prediction under rapidly changing wireless conditions is essential for aligning mobile networks with shifts in service demand. Traffic variations driven by crowd movement, large events, temporal usage patterns, and spatial–temporal interactions among BSs highlight the complexity of this problem. Some works concentrate primarily on improving prediction accuracy, whereas others develop broader operational frameworks that integrate predicted values into decision pipelines. Network traffic prediction approaches incorporate predictive outputs into operational decision-making. Some studies report operational mechanisms in which predicted load levels are used to adjust bandwidth power settings, or to reallocate resources (Bega et al., 2020; Hu et al., 2023; Rao et al., 2023; Tuna et al., 2024).

Traffic prediction research also includes applications in network slice dimensioning, where slice traffic predictions are used to determine slice-level resource allocation settings (Wu et al., 2022; Thantharate & Beard, 2022). Each network slice must be dynamically and securely managed for bandwidth, latency, and reliability, and be scalable to meet varying service requirements (Alwakeel & Alnaim, 2024). AI/ML-integrated network slicing management is a complex optimization problem, as slices are managed in real time according to changing load and service requirements (Huang et al., 2023; Lin et al., 2021).

Together, network traffic prediction and slice dimensioning illustrate how proactive, data-driven management is becoming essential in next-generation networks. Building on this direction, AI/ML enhances network performance and delivers a higher quality of service by enabling networks to perform operations such as traffic management, load balancing, and routing (Tshakwanda et al., 2024; Ding et al., 2025). For example, using RNN-based models, traffic density can be predicted using historical and current network data, allowing resources to be allocated in advance to areas expected to be congested, thereby preventing potential resource scarcity in the network.

Building on these AI/ML-driven network control examples, further studies explore intelligent network management mechanisms in greater detail. AI/ML has brought about a transformative change by providing dynamic, automatic, and intelligent management of networks with this capability. A recent study focuses on balancing the load across the network, particularly in dense urban areas, by enabling coordinated traffic management among different BSs and edge nodes in a decentralized manner (Jiang et al., 2024).

The AI/ML-based solution should be evaluated in terms of greener operations. A key research question is how 6G networks can support applications such as massive IoT, autonomous driving, and immersive VR while maintaining energy efficiency through intelligent resource allocation mechanisms (Zou et al., 2024). With AI/ML-based approaches, bandwidth and computing resources can be pre-allocated by predicting traffic density and user behavior, ensuring service quality during high-demand periods (Khan et al., 2022). Energy-aware AI models learn traffic patterns and place BSs into low-power modes during periods of low load (Alhussien & Aaron, 2024; Sami et al., 2021). Moreover, some studies also aim to reduce the energy footprint of the AI models themselves. Techniques such as model pruning, quantization, and lightweight architectures lower the computational cost of real-time inference, helping 6G move toward green communication (Alhussien & Aaron, 2024; I, 2021). Therefore, with 6G, AI/ML integration becomes not merely an enhancement but a necessity across operational, optimization, and design dimensions of mobile networks.

2.2. Toward AI-Native RAN: 3GPP and O-RAN Standardization Efforts

Only a subset of research proposals translates into standardization, which follows industry priorities and deployment needs. To move beyond, standards bodies such as 3GPP and O-RAN have been progressively defining structured frameworks for data collection, analytics, and operational intelligence. These efforts establish the technical foundations needed for scalable AI/ML in 6G.

Building on this framework, current studies within 3GPP examine enhancements to the air interface to better support AI/ML-based techniques (3GPP, 2025a). The goal is to define a common, reusable AI/ML architecture for the air interface that considers the impacts of performance, complexity, and standardization across different use cases. The framework specifies the principles for AI/ML models running at both the BS and UE, identifies which metrics or data structures require standardization, and outlines lifecycle procedures, including training, validation, inference, and updating. Within this scope, the evaluations focus on beam management, CSI feedback, and positioning accuracy enhancements. Moreover, 3GPP Release 20 exhibits a clear trend toward broadening AI/ML-driven 6G use cases, including low-overhead CSI/DMRS, SRS-based designs, PDCCH and HARQ enhancements, and RACH optimization (3GPP, 2025b).

Following the RAN Physical Layer evaluations, studies in the RAN Radio Protocol Layer 2 focus on how AI/ML can enhance radio-protocol behavior. In this context, the use cases addressed in (3GPP, 2025c) target connected-mode mobility, aiming to make measurement handling more efficient and cell-change procedures more predictable. The studies consider RRM measurement prediction, measurement event prediction and radio link failure (RLF) prediction. Across all these use cases, the common goal is to increase mobility stability-such as smoother cell switching and improved RLF management-while reducing measurement overhead and avoiding additional complexity on the network and UE sides.

Following the studies on the RAN Physical Layer and Layer 2, work in RAN Radio Protocol Layer 3 (RRC) examines how AI/ML can enhance link control, mobility, and configuration procedures. These efforts focus on making RRC-based functions more predictive and autonomous (3GPP, 2024a). The use cases are AI/ML-based coverage and capacity optimization (CCO), and AI/ML-powered slice optimization. Across both CCO and slice management, studies evaluate where training and inference should reside, what data needs to be shared, which additional information may be required on interfaces, and how AI/ML decisions can be integrated into existing RRC procedures.

Following the physical and protocol-level use cases at the RAN layers, system-level studies examine how AI/ML can be deployed, managed, and integrated across the end-to-end architecture. The first line of work focuses on the capabilities required to deploy, update, and migrate AI/ML models in a 5G system (3GPP, 2023). The objective is to balance the limited processing and storage resources of UEs with the higher computational power available in the network, establishing architectural requirements for transporting and operating AI/ML models. A second study examines architectural and functional extensions that allow the 5G system to support application-layer AI/ML operations (3GPP, 2022). This includes defining what information the network can expose, what conditions can be monitored, and which capabilities can be made available to the application layer. The final work defines the framework for managing AI/ML model lifecycle and inference functions in 5G systems (3GPP, 2025d).

As we approach the 6G era, the telecommunications community positions the RAN as an AI-native system that inherently incorporates data collection, model execution, model transfer, and continuous learning (3GPP, 2024b). Realizing this vision requires a unified architecture that integrates both AI for network-where in-network functions are driven by AI-and network for AI-where the network supports third-party AI/ML operations (3GPP, 2025e). This leads to a distributed intelligence model in which the UE and BS execute and update models cooperatively in real time, enabling the air interface to adapt dynamically to environmental conditions, load variations, interference patterns, and mobility behaviors (3GPP, 2025f).

Across ongoing 6G discussions, there is broad alignment that the first 6G release will require a common AI/ML lifecycle management framework capable of supporting all major use cases. This consensus also emphasizes the need for a modular, extensible, and continuous learning AI-native architecture that remains future-proof as new capabilities emerge (3GPP, 2025g). Moreover, they emphasize that transforming the network architecture holistically into a data-driven, programmable, and self-adaptive architecture. As a natural extension of this vision, the O-RAN ecosystem has defined a native AI approach to prepare the open RAN architecture for 6G and developed a comprehensive architectural framework for the embedded and distributed operation of AI across the mobile network (O-RAN, 2023a). In this approach, data collection, model training, distribution, inference, and update cycles are considered inherent components of the RAN architecture, with AI functions intended to be embedded across all layers (O-RAN, 2023b).

In conclusion, O-RAN's native AI architecture and the evolution of Near-RT RIC for 6G provide a strategic framework that complements and operationally implements the AI-native RAN vision put forth by 3GPP and industry players. This framework aims to preserve the open, modular, and multi-vendor RAN structure in the 6G era and to bring together the programmability, data access, distributed intelligence, and continuous learning mechanisms required by AI/ML-based radio design under a holistic architecture (O-RAN, 2025).

3. Challenges Toward AI-Native 6G Networks

Integrating AI/ML into next-generation wireless networks introduces a unique blend of scientific, architectural, and operational challenges. Unlike other application domains, wireless environments exhibit high levels of mobility, interference, fading, hardware impairments, and scheduling variability. These factors make the underlying data distribution highly non-stationary and often poorly labeled. Because radio conditions change quickly, models trained on past data can become inaccurate in a short time, which has increased interest in learning methods that can adapt during operation. At the same time, wireless networks rarely offer clean or frequent labels; important events are uncommon, and performance-related information often arrives with some delay. As a result, the learning process becomes more challenging, motivating the use of semi-supervised, self-supervised, and weakly supervised approaches that can effectively handle noisy or incomplete measurements. In addition, strict PHY/MAC timing constraints-often on the order of sub-milliseconds-necessitate lightweight architectures, efficient feature design, and low-precision inference to make AI/ML deployment feasible at scale.

Beyond these scientific and data-driven issues, achieving robust generalization across diverse environments remains a major research frontier. AI/ML models must operate reliably across a wide range of antenna setups, coverage conditions, mobility levels, and traffic patterns, ideally without requiring retraining for every new scenario. This expectation becomes even more crucial in multi-BS architectures, particularly in environments where non-terrestrial and terrestrial networks coexist, and in the diverse settings anticipated for 5G-Advanced and 6G. Privacy and security add another layer of challenge, since mobility traces, channel characteristics, and traffic behavior can expose sensitive information. For this reason, learning systems must be robust against risks such as data leakage, adversarial actions, and various types of poisoning. These challenges have driven the field toward approaches that protect user data, support distributed or federated training, and remain vigilant against potential attacks-while still meeting performance expectations and regulatory requirements.

From a deployment standpoint, introducing AI/ML into commercial networks presents system-level challenges that extend beyond improving model accuracy. Distributed RAN setups, the use of multiple frequency layers, non-terrestrial and aerial platforms, and multi-vendor environments can all lead to differences in how features are collected, how measurements are processed, and at what granularity they are available. When the location of the data and the point where inference is performed are not aligned, additional issues arise-most notably delays and limited visibility. These effects are especially critical for functions like beam management, CSI reporting, mobility decisions, and interference prediction, where the required reaction times are tightly linked to protocol timing. Ensuring multi-vendor interoperability, therefore, requires standardized feature structures, unified model interfaces, and coordination between 3GPP and O-RAN frameworks to avoid fragmented AI/ML ecosystems.

Finally, lifecycle management and system reliability define another critical research frontier. Commercial networks must manage the full lifecycle of models-training, validation, distribution, activation, rollback, and continuous monitoring-across thousands of cells and heterogeneous hardware platforms. Energy and computation budgets at UEs, DUs, CUs, and edge cloud systems constrain where and how AI/ML workloads can be executed. In addition, model files, inference outputs, and intermediate representations can introduce new points of vulnerability, which makes security and integrity essential for maintaining stable network operation. Taken together, these challenges demonstrate that AI-native 6G will necessitate solutions that integrate algorithms, architectures, protocols, and operational practices-rather than relying solely on improved model performance.

Conclusion

The shift toward AI-nativeness marks a significant transformation in how modern mobile networks are designed, optimized, and managed. The evolution toward IMT-2030 has elevated intelligence to a core architectural principle. Growing service heterogeneity and rising traffic demands, together with increasing RAN complexity, are positioning AI/ML as an essential enabler of predictive, autonomous, and energy-efficient operation.

Numerous research studies show how AI/ML-enabled methods can significantly improve responsiveness and efficiency. Yet turning these advances into deployable solutions requires structured development pipelines, system-level validation, and alignment with standardized interfaces. This need has driven 3GPP and O-RAN toward the development of coordinated frameworks for data collection, model training, inference, and lifecycle management.

The progress toward AI-native 6G will require tightly integrated algorithms, interfaces, protocols, and operational models. Standardization is already evolving in this direction, with unified AI/ML lifecycle management, distributed learning across UE and network nodes, and mechanisms that embed intelligence directly into the air interface and end-to-end architecture. In parallel, industry efforts enhance openness, programmability, and multi-vendor interoperability, which are crucial for a sustainable AI-native RAN design.

Despite significant progress, key research frontiers remain open, including robust generalization, privacy-preserving learning, ultra-low-latency inference, cross-layer coordination, and energy-efficient edge AI. Addressing these challenges will require sustained and coordinated efforts across academia, industry, and standardization bodies. Taken together, these developments highlight the growing role of AI/ML as the basis for connectivity that is inherently adaptive and intelligence-driven. As mobile networks advance toward 6G, success will depend on bridging scientific insight with architectural innovation and operational readiness, enabling RANs to learn continuously, adapt autonomously, and maintain intelligent performance at scale.

References

Alhussien, N., & Aaron, G. T. (2024). Toward AI-enabled green 6G networks: A resource management perspective. *IEEE Access, 12*, 132972–132995. https://doi.org/10.1109/ACCESS.2024.3460656

Alwakeel, A., & Alnaim, A. (2024). Network slicing in 6G: A strategic framework for IoT in smart cities. *Sensors, 24*(13), 4254. https://doi.org/10.3390/s24134254

Balmer, R., Levin, S., & Schmidt, S. (2020). Artificial intelligence applications in telecommunications and other network industries. *Telecommunications Policy, 44*(6), 101977. https://doi.org/10.1016/j.telpol.2020.101977

Bega, D., Gramaglia, M., Fiore, M., Banchs, A., & Costa-Pérez, X. (2020). DeepCog: Optimizing resource provisioning in network slicing with AI-based capacity forecasting. *IEEE Journal on Selected Areas in Communications, 38*(2), 361–376. https://doi.org/10.1109/JSAC.2019.2959245

Brilhante, D., Manjarres, J. C., Moreira, R., de Oliveira Veiga, L., de Rezende, J., Müller, F., Klautau, A., Mendes, L. L., & de Figueiredo, F. P. (2023). A literature survey on AI-aided beamforming and beam management for 5G and 6G systems. *Sensors, 23*(9), 4359. https://doi.org/10.3390/s23094359

Charan, G., Alrabeiah, M., & Alkhateeb, A. (2021). Vision-aided 6G wireless communications: Blockage prediction and proactive handoff. *IEEE Transactions on Vehicular Technology, 70*(10), 10193–10208. https://doi.org/10.1109/TVT.2021.3104219

Chataut, R., Nankya, M., & Akl, R. (2024). 6G networks and the AI revolution—Exploring technologies, applications, and emerging challenges. *Sensors, 24*(6), 1888. https://doi.org/10.3390/s24061888

Ding, C., Zhu, L., Shen, L., Li, Z., Li, Y., & Liang, Q. (2025). The intelligent traffic flow control system based on 6G and optimized genetic algorithm. *IEEE Transactions on Intelligent Transportation Systems, 26*(7), 17403–17416. https://doi.org/10.1109/TITS.2024.3467269

Dulaj, K., Alhammadi, A., Shayea, I., El-Saleh, A., & Alnakhli, M. (2025). Harnessing machine learning for intelligent networking in 5G technology and beyond: Advancements, applications and challenges. *IEEE Open Journal of Intelligent Transportation Systems, 6*, 605–633. https://doi.org/10.1109/OJITS.2025.3564361

García, C. R., Bouchmal, O., Stan, C., Giannakopoulos, P., Cimoli, B., Olmos, J. J., Rommel, S., & Monroy, I. T. (2023). Secure and agile 6G networking—Quantum and AI enabling technologies. In *2023 23rd International Conference on Transparent Optical Networks (ICTON)* (pp. 1–4). IEEE. https://doi.org/10.1109/ICTON59386.2023.10207418

Hu, Y., Zhou, Y., Song, J., Xu, L., & Zhou, X. (2023). Citywide mobile traffic forecasting using spatial-temporal downsampling transformer neural networks. *IEEE Transactions on Network and Service Management, 20*(1), 152–165. https://doi.org/10.1109/TNSM.2022.3214483

Huang, J., Yang, F., Chakraborty, C., Guo, Z., Zhang, H., Zhen, L., & Yu, K. (2023). Opportunistic capacity based resource allocation for 6G wireless systems with network slicing. *Future Generation Computer Systems, 140*, 390–401. https://doi.org/10.1016/j.future.2022.10.032

I, C.-L. (2021). AI as an essential element of a green 6G. *IEEE Transactions on Green Communications and Networking, 5*(1), 1–3. https://doi.org/10.1109/TGCN.2021.3057247

ITU-R. (2008). *Requirements related to technical performance for IMT-Advanced radio interface(s) (Recommendation ITU-R M.2134-0).* International Telecommunication Union.

ITU-R. (2015). *IMT vision—Framework and overall objectives of the future development of IMT for 2020 and beyond (Recommendation ITU-R M.2083-0).* International Telecommunication Union.

ITU-R. (2023). *Framework and overall objectives of the future development of IMT for 2030 and beyond (Recommendation ITU-R M.2160-0).* International Telecommunication Union.

Jiang, F., Peng, Y., Dong, L., Wang, K., Yang, K., Pan, C., Niyato, D., & Dobre, O. (2024). Large language model enhanced multi-agent systems for 6G communications. *IEEE Wireless Communications, 31*(6), 48–55. https://doi.org/10.1109/MWC.016.2300600

Jiao, L., Shao, Y., Sun, L., Liu, F., Yang, S., Ma, W., Li, L., Liu, X., Hou, B., Zhang, X., Shang, R., Li, Y., Wang, S., Tang, X., & Guo, Y. (2024). Advanced deep learning models for 6G: Overview, opportunities, and challenges. *IEEE Access, 12*, 133245–133314. https://doi.org/10.1109/ACCESS.2024.3418900

Kaur, J., Khan, M., Iftikhar, M., Imran, M., & Emad Ul Haq, Q. (2021). Machine learning techniques for 5G and beyond. *IEEE Access, 9*, 23472–23488. https://doi.org/10.1109/ACCESS.2021.3051557

Khan, S., Hussain, A., Nazir, S., Khan, F., Oad, A., & Alshehri, M. D. (2022). Efficient and reliable hybrid deep learning-enabled model for congestion control in 5G/6G networks. *Computer Communications, 182*, 31–40. https://doi.org/10.1016/j.comcom.2021.11.001

Lin, K., Li, Y., Zhang, Q., & Fortino, G. (2021). AI-driven collaborative resource allocation for task execution in 6G-enabled massive IoT. *IEEE Internet of Things Journal, 8*(7), 5264–5273. https://doi.org/10.1109/JIOT.2021.3051031

Mahmood, M., Matin, M. A., Sarigiannidis, P., & Goudos, S. (2022). A comprehensive review on artificial intelligence/machine learning algorithms for empowering the future IoT toward 6G era. *IEEE Access, 10*, 87535–87562. https://doi.org/10.1109/ACCESS.2022.3199689

O-RAN ALLIANCE. (2023a). *O-RAN native AI architecture description* (RR-2023-02).

O-RAN ALLIANCE. (2023b). *Research report on native and cross-domain AI: State of the art and future outlook* (RR-2023-03).

O-RAN ALLIANCE. (2025). *Evolution of O-RAN Near-RT RIC toward 6G* (RR-2025-04).

Rao, Z., Xu, Y., Pan, S., Guo, J., Yan, Y., & Wang, Z. (2023). Cellular traffic prediction: A deep learning method considering dynamic nonlocal spatial correlation, self-attention, and correlation of spatiotemporal feature fusion. *IEEE Transactions on Network and Service Management, 20*(1), 426–440. https://doi.org/10.1109/TNSM.2022.3187251

Rashid, A., & Kausik, M. (2024). AI revolutionizing industries worldwide: A comprehensive overview of its diverse applications. *Hybrid Advances, 7*, 100277. https://doi.org/10.1016/j.hybadv.2024.100277

Sami, H., Otrok, H., Bentahar, J., & Mourad, A. (2021). AI-based resource provisioning of IoE services in 6G: A deep reinforcement learning approach. *IEEE Transactions on Network and Service Management, 18*(3), 3527–3540. https://doi.org/10.1109/TNSM.2021.3066625

Sanjalawe, Y., Fraihat, S., Al-E'Mari, S., Abualhaj, M., Makhadmeh, S., & Alzubi, E. (2025). A review of 6G and AI convergence: Enhancing communication networks with artificial intelligence. *IEEE Open Journal of the Communications Society, 6*, 2308–2355. https://doi.org/10.1109/OJCOMS.2025.3553302

Shah, B. M., Murtaza, M., & Raza, M. (2020). Comparison of 4G and 5G cellular network architecture and proposing of 6G, a new era of AI. In *2020 5th International Conference on Innovative Technologies in Intelligent Systems and Industrial Applications (CITISIA)*. IEEE. https://doi.org/10.1109/CITISIA50690.2020.9371846

Shi, Y., Lian, L., Shi, Y., Wang, Z., Zhou, Y., Fu, L., Bai, L., Zhang, J., & Zhang, W. (2023). Machine learning for large-scale optimization in 6G wireless networks. *IEEE Communications Surveys & Tutorials, 25*(3), 2088–2132. https://doi.org/10.1109/COMST.2023.3300664

Taleb, T., Benzaïd, C., Addad, R. A., & Samdanis, K. (2023). AI/ML for beyond 5G systems: Concepts, technology enablers & solutions. *Computer Networks, 237*, 110044. https://doi.org/10.1016/j.comnet.2023.110044

Thantharate, A., & Beard, C. (2023). ADAPTIVE6G: Adaptive resource management for network slicing architectures in current 5G and future 6G systems. *Journal of Network and Systems Management, 31*, Article 9. https://doi.org/10.1007/s10922-022-09693-1

Tshakwanda, P. M., Arzo, S. T., & Devetsikiotis, M. (2024). Advancing 6G network performance: AI/ML framework for proactive management and dynamic optimal routing. *IEEE Open Journal of the Computer Society, 5*, 303–314. https://doi.org/10.1109/OJCS.2024.3398540

Tuna, E., Baykal, A., & Soysal, A. (2024). Multivariate and multistep mobile traffic prediction with SLA constraints: A comparative study. *Ad Hoc Networks, 163*, 103594. https://doi.org/10.1016/j.adhoc.2024.103594

Usman, Y., Oladipupo, H., During, A., Akl, R., & Chataut, R. (2025). AI, ML, and LLM integration in 5G/6G networks: A comprehensive survey of architectures, challenges, and future directions. *IEEE Access, 13*, 168914–168950. https://doi.org/10.1109/ACCESS.2025.3608736

Wang, J., Liu, J., Li, J., & Kato, N. (2023). Artificial intelligence-assisted network slicing: Network assurance and service provisioning in 6G. *IEEE Vehicular Technology Magazine, 18*(1), 49–58. https://doi.org/10.1109/MVT.2022.3228399

Wu, W., Zhou, C., Li, M., Wu, H., Zhou, H., Zhang, N., Shen, X. S., & Zhuang, W. (2022). AI-native network slicing for 6G networks. *IEEE Wireless Communications, 29*(1), 96–103. https://doi.org/10.1109/MWC.001.2100338

Zawish, M., Dharejo, F. A., Khowaja, S. A., Raza, S., Davy, S., Dev, K., & Bellavista, P. (2024). AI and 6G into the metaverse: Fundamentals, challenges and future research trends. *IEEE Open Journal of the Communications Society, 5*, 730–778. https://doi.org/10.1109/OJCOMS.2024.3349465

Zhang, C., Patras, P., & Haddadi, H. (2019). Deep learning in mobile and wireless networking: A survey. *IEEE Communications Surveys & Tutorials, 21*(3), 2224–2287. https://doi.org/10.1109/COMST.2019.2904897

Zhang, C., Ueng, Y.-L., Studer, C., & Burg, A. (2020). Artificial intelligence for 5G and beyond 5G: Implementations, algorithms, and optimizations. *IEEE Journal on Emerging and Selected Topics in Circuits and Systems, 10*(2), 149–163. https://doi.org/10.1109/JETCAS.2020.3000103

Zhao, J., Liu, C., Liao, J., & Wang, D. (2024). Deep learning in wireless communications for physical layer. *Physical Communication, 67*, 102503. https://doi.org/10.1016/j.phycom.2024.102503

Zou, S., Wu, J., Yu, H., Wang, W., Huang, L., Ni, W., & Liu, Y. (2024). Efficiency-optimized 6G: A virtual network resource orchestration strategy by enhanced particle swarm optimization. *Digital Communications and Networks, 10*(5), 1221–1233. https://doi.org/10.1016/j.dcan.2023.06.008

Zuo, Y., Guo, J., Gao, N., Zhu, Y., Jin, S., & Li, X. (2023). A survey of blockchain and artificial intelligence for 6G wireless communications. *IEEE Communications Surveys & Tutorials, 25*(4), 2494–2528. https://doi.org/10.1109/COMST.2023.3315374

3GPP. (2022). *TR 23.700-80 study on 5G system support for AI/ML-based services (Release 18).*

3GPP. (2023). *TR 22.876 study on AI/ML model transfer Phase 2 (Release 19).*

3GPP. (2024a). *TR 38.743 study on enhancements for AI/ML for NG-RAN.*

3GPP. (2024b). *6GWS-250102 Intel views on 6G RAN and radio interface.*

3GPP. (2025a). *TR 38.843 study on AI/ML for NR air interface (Release 19).*

3GPP. (2025b). *R1-2509598 moderator summary #5 on AI/ML for 6GR.*

3GPP. (2025c). *3GPP TR 38.744 study on AI/ML for mobility in NR (Release 19).*

3GPP. (2025d). *TS 28.105 AI/ML management (Release 19).*

3GPP. (2025e). *6GWS-250159 views on 6G radio.*

3GPP. (2025f). *6GWS-250163 vision and priorities for next generation radio technology.*

3GPP. (2025g). *6GWS-250004 6G radio and RAN.*

About the Authors

Dr. Evren TUNA

Ulak Communications | evren.tuna[at]ulakhaberlesme.com.tr
ORCID: 0000-0001-6551-6460

Dr. Evren Tuna received his Ph.D. in Electrical and Electronics Engineering from Bahçeşehir University and his M.Sc. and B.Sc. degrees in Electronics and Communication Engineering from İzmir Institute of Technology. He is a Lead R&D Engineer at Ulak Communications and Head of the UL6B – ULAK 6G and Beyond Communication Technologies Laboratory, established under the TÜBİTAK 1515 Frontier R&D Laboratory Support Programme, where he leads research on 5G-Advanced and 6G technologies. He serves as the 3GPP contact person and represents Istanbul Medipol University as an active delegate in RAN1 and RAN2. He has co-authored multiple international publications and holds several patents in wireless communications. He teaches Digital Communications at Istanbul Okan University and has previously taught at Yıldız Technical University. His research interests include 5G-Advanced and 6G systems, AI/ML-driven network intelligence, mobility management, NTN, channel modeling, and data-driven traffic prediction.

Prof. Dr. Hüseyin ARSLAN

Istanbul Medipol University | huseyinarslan[at]medipol.edu.tr
ORCID: 0000-0001-9474-7372

Hüseyin Arslan (IEEE Fellow, NAI Fellow, IEEE Distinguished Lecturer, Member of the Turkish Academy of Sciences) received his B.S. degree from Middle East Technical University and his M.S. and Ph.D. degrees from Southern Methodist University. He served as a Professor at the University of South Florida and has been the Founding Dean of the School of Engineering and Natural Sciences at Istanbul Medipol University since 2013. Throughout his career, he has worked closely with academia and industry through research collaborations and consultancy roles with Ericsson, Anritsu, Savronik, TÜBİTAK, and ULAK Communications, and previously served on the Board of Directors of Turkcell. His research focuses on wireless communication systems, particularly PHY and MAC layer design. He has authored nearly 600 peer-reviewed journal and conference papers, contributed to 63 book chapters, and holds 84 issued and over 80 pending patents. His work has received more than 26,400 citations (h-index: 66) and reflects a sustained balance between theoretical research and practical implementation.