



AI-ENABLED 6G NETWORKS FOR INTELLIGENT BROADBAND WIRELESS COMMUNICATIONS

Dr. P. Anbarasi

*Assistant Professor & Head,
Department of Computer Science,
Sri Muthukumaran Arts and Science College,
Chennai, Tamil Nadu, India.*

Mrs. Rajapriya C

*Assistant Professor,
Department of Computer Science,
Sri Muthukumaran Arts and Science College,
Chennai, Tamil Nadu, India.*

Abstract:

The sixth generation (6G) of wireless communication systems is envisioned to deliver terabit-per-second data rates, sub-millisecond latency, ubiquitous three-dimensional coverage, and native support for massive machine-type and holographic communications. Achieving these ambitious targets requires networks that can sense, learn, reason, and act autonomously, which places Artificial Intelligence (AI) and Machine Learning (ML) at the very core of the 6G design philosophy rather than as an add-on feature, as was largely the case in earlier generations. This paper presents a comprehensive study of AI-enabled 6G networks for intelligent broadband wireless communication. We first review the key performance requirements and architectural evolution from 5G to 6G, and then examine how AI/ML techniques—including deep learning, federated learning, reinforcement learning, and large AI models—are applied to spectrum management, intelligent network slicing, massive multiple-input multiple-output (MIMO) beamforming, self-organizing networks, and physical-layer security. We

further discuss key enabling technologies such as terahertz (THz) communication, reconfigurable intelligent surfaces (RIS), integrated sensing and communication (ISAC), digital twin networks, and Open RAN, and describe how their convergence with AI forms a coherent AI-native 6G ecosystem. A layered AI-enabled 6G framework is proposed, and an illustrative system-level simulation study is presented to demonstrate the potential gains of AI-driven resource management over conventional heuristic approaches in terms of throughput, latency, and energy efficiency. Finally, open research challenges relating to explainability, trustworthiness, energy sustainability, standardization, and security are identified, along with promising future research directions. The study aims to serve as a consolidated reference for researchers working at the intersection of artificial intelligence and next-generation wireless communication systems.

Keywords: 6G networks, artificial intelligence, machine learning, deep learning, federated learning, network slicing, massive MIMO, reconfigurable intelligent surfaces, terahertz communication, integrated sensing and communication, digital twin, Open RAN.

I. Introduction

Over the past four decades, cellular communication systems have evolved through successive generations, each redefining the boundary of what wireless networks could deliver. The fourth generation (4G) introduced mobile broadband, and the fifth generation (5G) extended this with enhanced mobile broadband (eMBB), ultra-reliable low-latency communication (URLLC), and massive machine-type communication (mMTC). Despite these advances, 5G networks still rely heavily on static, rule-based optimization and human-in-the-loop network planning, which limits their ability to cope with highly dynamic traffic patterns, heterogeneous device populations, and increasingly complex service requirements.

The sixth generation (6G) of wireless networks, targeted for commercial deployment around 2030, is expected to support peak data rates on the order of 1 Tb/s, end-to-end latency below 0.1 ms, ultra-massive connectivity densities, and integrated sensing, computing, and communication capabilities. Meeting such requirements with conventional network management techniques is impractical because the underlying optimization problems are non-convex, high-dimensional, and must be solved

in near real time across highly heterogeneous infrastructure spanning terrestrial, aerial, and satellite segments.

Artificial Intelligence (AI) has therefore emerged as a foundational pillar of 6G rather than a peripheral enhancement. AI-native network design proposes that intelligence be embedded from the physical layer to the application layer, enabling the network to autonomously sense its environment, predict traffic and channel behavior, allocate resources, detect anomalies, and heal itself with minimal human intervention [1], [2]. This paradigm shift is often referred to as "AI-native 6G" or "intelligent 6G," distinguishing it from earlier efforts that merely bolted machine learning modules onto existing 4G/5G architectures [3].

This paper provides a structured overview of how AI and 6G converge to enable intelligent broadband wireless communication. The main contributions are as follows: (i) a review of 6G requirements and architectural evolution; (ii) a systematic discussion of AI/ML applications across the 6G protocol stack; (iii) an overview of enabling technologies including THz communication, RIS, ISAC, digital twins, and Open RAN; (iv) a proposed layered AI-enabled 6G framework; (v) an illustrative performance evaluation comparing AI-driven and conventional resource management; and (vi) a discussion of open challenges and future research directions.

This survey works through the The remainder of this paper is organized as follows. Section II reviews related work. Section III describes 6G requirements and network architecture. Section IV discusses the role of AI in 6G networks. Section V presents key enabling technologies. Section VI formulates a representative AI-driven resource-allocation problem. Section VII presents representative use cases. Section VIII proposes an AI-enabled 6G framework. Section IX presents an illustrative performance evaluation. Section X discusses standardization efforts. Section XI discusses open challenges, Section XII outlines future research directions, and Section XIII concludes the paper.

II. Related Work

A growing body of literature has examined the role of AI in future wireless networks. Yang et al. [2] provided one of the earliest comprehensive treatments of artificial-intelligence-enabled intelligent 6G networks, outlining how deep learning and reinforcement learning could be embedded across the network stack. Jiang et al. [3] surveyed machine learning paradigms – supervised, unsupervised, and reinforcement learning – for next-generation wireless networks, establishing a taxonomy that later works have extended to 6G. Du et al. [6] examined how machine learning can enhance bandwidth utilization, massive access, and ultra-reliable low-latency services in 6G systems.

More recent surveys have focused on specific sub-domains. Shi et al. [1] reviewed machine learning techniques for large-scale, non-convex optimization problems that arise in 6G resource management. Bonati et al. [11] and Niknam et al. [4] examined intelligence embedded in Open RAN (O-RAN) architectures, while Giannopoulos et al. [5] discussed AI/ML workflows for disaggregated radio access networks. Federated learning has received particular attention as a privacy-preserving distributed learning paradigm suited to 6G; Hasan et al. [10] reviewed its requirements and applications for 6G communication. Sheraz et al. [18] surveyed the emerging concept of AI-enabled digital twin networks for 6G, and Clerckx et al. [20] provided a tutorial on multiple access techniques for intelligent and multifunctional 6G systems.

Physical-layer and air-interface aspects of AI-enabled 6G have also received substantial attention. Guan et al. [12] examined AI-based efficiency evaluation for massive MIMO systems operating on high-altitude platform stations, while Salh et al. [13] and Liu et al. [14] surveyed deep learning approaches for meeting ultra-reliable, low-latency requirements at the physical and MAC layers. At a broader ecosystem level, De Alwis et al. [16], Quy et al. [17], and Tera et al. [18] presented comprehensive surveys of 6G architecture, applications, and enabling technologies, situating AI as one of several converging trends alongside THz communication, non-terrestrial networks, and blockchain-based trust management.

Agiwal et al. [21] and Bhat and Alqahtani [22] additionally provide useful historical context by tracing the transition from 5G requirements to the broader 6G ecosystem vision. Explainability and trustworthiness of AI in 6G have also emerged as active research themes. Fiandrino et al. [7] and Khan et al. [8] highlighted the need for native explainable and robust AI to ensure that autonomous network decisions remain interpretable and auditable. Chen et al. [15] discussed the opportunities and challenges of deploying large AI models—including large language models and foundation models—within 6G networks, and Jiang et al. [24] proposed multi-agent systems enhanced by large language models for autonomous network communications and orchestration.

Alhussien and Gulliver [19] further examined AI-enabled resource management from an energy-sustainability perspective, underscoring the growing recognition that the computational cost of AI itself must be accounted for in 6G's overall energy budget.

Table V positions the present paper relative to representative prior surveys. Whereas most existing works concentrate on a single sub-domain—such as resource optimization [1], Open RAN intelligence [4], [5], [11], federated learning [10], or digital twins [18]—this paper aims to connect these strands into a single, layered AI-enabled 6G framework, and complements the conceptual discussion with a proposed system model and an illustrative quantitative evaluation.

Table V: Positioning of This Paper Relative to Representative Prior Surveys

Reference	Primary Focus	Framework Proposed	Quantitative Evaluation?
Shi et al. [1]	Large-scale ML optimization	No	Survey only
Bonati et al. [11]	O-RAN intelligence	Partial (RIC)	No
Hasan et al. [10]	Federated learning for 6G	No	No
Sheraz et al. [18]	AI-enabled digital twins	No	No
This paper	AI across full 6G stack	Yes (4-layer)	Yes (illustrative)

III. 6G Network Requirements and Architecture

A. Key Performance Requirements

6G networks are designed to substantially exceed the capabilities of 5G across several performance dimensions. Table I summarizes representative target requirements commonly cited in the 6G literature [16], [17], [21], [22], relative to typical 5G specifications.

Table I: Representative 5G vs. 6G Performance Targets

Metric	5G (typical)	6G (target)
Peak data rate	20 Gb/s	≥ 1 Tb/s
User-plane latency	1 ms	< 0.1 ms
Connection density	10^6 devices/km ²	10^7 – 10^8 devices/km ²
Mobility support	500 km/h	> 1000 km/h
Reliability	99.999%	99.99999%

Coverage	Terrestrial-centric	Space-air-ground-sea
Positioning accuracy	~10 m	~1 cm (indoor/outdoor)
Energy efficiency	Baseline	10-100× improvement

Beyond raw performance figures, 6G introduces qualitatively new capabilities such as native sensing (ISAC), pervasive AI inference at the edge, holographic and extended-reality communication, and support for digital twins of both the network and the physical world it serves [17].

B. Traffic and Service Diversity

In addition to the quantitative targets in Table I, 6G must accommodate a substantially more diverse service mix than 5G. Beyond the three canonical 5G service classes, 6G is expected to natively support new categories including holographic-type communication, extremely reliable and low-latency communication for tactile/haptic applications, mobile broadband reliable low-latency communication combining high throughput with strict latency bounds, and massive-scale integrated sensing services that were not part of the original IMT-2020 framework. Meeting the combined resource, latency, and reliability requirements of these overlapping service classes on a shared physical infrastructure is one of the primary motivations for AI-driven, rather than static, resource orchestration.

C. Architectural Evolution

The architectural evolution toward 6G is characterized by three broad shifts. First, a

move from terrestrial-only networks to integrated space-air-ground-sea networks (SAGSIN), combining terrestrial base stations with high- and low-altitude platforms and dense satellite constellations to guarantee ubiquitous global coverage.

Second, a move from hardware-defined to fully software-defined, virtualized, and disaggregated infrastructure, exemplified by the continued evolution of Open RAN, network function virtualization (NFV), and cloud-native cores [4], [5], [11]. Third, and most importantly for this paper, a move from AI-assisted to AI-native design, in which learning-based control loops are embedded as first-class citizens of the architecture at every layer, from the radio access network (RAN) to orchestration and management planes.

This AI-native philosophy is often realized through a closed-loop "sense-learn-decide-act" cycle: the network continuously senses radio and traffic conditions, learns patterns using onboard or edge-hosted AI models, decides on control actions (e.g., beam selection, slice reconfiguration, power allocation), and acts on the infrastructure – all with minimal latency and human involvement. Standardization bodies including 3GPP and the O-RAN Alliance have begun incorporating AI/ML workflows into RAN intelligent controllers (RIC), reflecting this shift toward pervasive, standardized network intelligence [9], [11].

D. Non-Terrestrial and Space-Air-Ground Integration

A distinguishing feature of 6G relative to earlier generations is the native integration of non-terrestrial networks—low Earth orbit (LEO) satellite constellations, high-altitude platform stations (HAPS), and unmanned aerial vehicles—alongside terrestrial base stations to guarantee global, ubiquitous coverage, including over oceans, remote regions, and disaster-affected areas where terrestrial infrastructure is unavailable [17], [22]. Coordinating handovers and resource allocation across such a heterogeneous, highly dynamic multi-layer topology, where satellite links exhibit rapidly time-varying delay and Doppler characteristics, is again a task for which AI-based prediction and control is considered essential rather than optional, since the underlying topology and channel conditions change far more rapidly than in fixed terrestrial deployments.

IV. Role of Artificial Intelligence in 6G Networks

AI techniques are being embedded across virtually every functional block of the 6G network. This section discusses six representative application domains: spectrum management, network slicing, massive MIMO/beamforming, self-organizing networks, federated/edge learning, and network security. Table VI maps the dominant AI/ML technique families to these application domains and summarizes the primary benefit each brings, before the remainder of this section discusses each domain in detail.

Table VI: AI/ML Technique Families Mapped to 6G Application Domains

Application Domain	Dominant AI/ML Techniques	Primary Benefit
Spectrum management	CNN/RNN sensing, multi-agent RL	Real-time interference-aware access
Network slicing	LSTM forecasting, hierarchical DRL	Proactive, SLA-aware orchestration
Massive MIMO/beamforming	Autoencoders, supervised DL	Reduced pilot/feedback overhead
Self-organizing networks	Anomaly detection, DRL	Autonomous healing & optimization
Federated/edge learning	Federated averaging, split learning	Privacy-preserving distributed AI
Network security	DL classifiers, adversarial training	Real-time intrusion/anomaly detection

A. AI-Driven Spectrum Sensing and Management

6G networks will operate across an extremely wide range of frequencies, from sub-6 GHz bands to millimeter-wave and terahertz spectrum, each with distinct propagation characteristics. Deep learning models, particularly convolutional and recurrent neural networks, are used for spectrum sensing, interference prediction, and dynamic spectrum access, allowing the network to identify unused spectrum holes and reallocate them in real time [1], [6]. Reinforcement learning (RL) agents can further learn near-optimal channel access and power control policies directly from

interaction with the radio environment, outperforming static allocation schemes under non-stationary traffic conditions.

A particularly important use case is cognitive spectrum sharing among heterogeneous radio access technologies co-existing within the same geographic area. Multi-agent reinforcement learning (MARL) allows distributed base stations and access points to learn cooperative or competitive spectrum access policies without requiring a centralized controller to solve a combinatorial assignment problem for every scheduling interval.

Graph neural networks (GNNs) have also been explored to model the topological relationships among interfering links, enabling the network to generalize spectrum management policies to previously unseen cell layouts and user distributions [1], [6].

B. Intelligent Network Slicing and Orchestration

Network slicing partitions a shared physical infrastructure into multiple logical networks, each tailored to a specific service (e.g., eMBB, URLLC, mMTC). Because slice demand fluctuates continuously, AI-based forecasting models (e.g., long short-term memory networks) are used to predict traffic loads and proactively scale slice resources, while deep reinforcement learning (DRL) agents handle admission control, resource orchestration, and inter-slice isolation to maximize utilization while meeting per-slice service-level agreements [1], [4].

Beyond intra-domain slicing, 6G is expected to support end-to-end slices that span radio access, transport, and core network domains, as well as multiple administrative operators. Hierarchical DRL architectures—where a high-level agent allocates coarse-grained resources across domains and low-level agents optimize fine-grained scheduling within each domain—have been proposed to make this cross-domain orchestration problem tractable.

Digital-twin-assisted slicing, discussed further in Section V-D, additionally allows candidate slice configurations to be validated in simulation prior to activation, reducing the risk of SLA violations during live reconfiguration.

C. AI-Based Massive MIMO and Beamforming

Massive MIMO and its evolution toward extremely large-scale antenna arrays are central to achieving 6G's spectral efficiency targets, but channel estimation and beamforming become computationally prohibitive as the number of antennas grows. Deep learning-based channel estimators, autoencoders for channel state information (CSI) compression and feedback, and learned beamforming policies significantly reduce pilot overhead and computational complexity compared with conventional model-based approaches [12], [21]. AI-assisted hybrid beamforming further enables joint optimization of analog and digital precoders under hardware constraints.

For extremely large aperture arrays, near-field propagation effects become significant, invalidating the far-field planar-wave assumptions used in conventional beamforming algorithms. Deep learning models trained on near-field channel data can implicitly capture spherical-wave characteristics and produce beam patterns that account for both angle and distance, a capability that is difficult to achieve analytically at scale.

In parallel, beam prediction models that fuse multi-modal side information—such as user location, camera, and LiDAR data available in ISAC-enabled deployments—can proactively select beams before a handover event occurs, reducing beam-training overhead and mobility-induced outages [12].

D. Self-Organizing Networks and Autonomous Operations

Self-organizing network (SON) functions—self-configuration, self-optimization, and self-healing—are essential for managing the scale and heterogeneity of 6G deployments. AI models continuously monitor key performance indicators (KPIs) to detect anomalies such as cell outages or coverage holes, and trigger automated remediation actions such as parameter reconfiguration or traffic offloading, reducing operational expenditure and improving reliability [7], [8].

The maturity of autonomous network operation is often described using a staged autonomy model, progressing from AI-assisted troubleshooting (where a human

operator still approves recommended actions), to conditional autonomy (where routine actions are executed automatically but complex incidents are escalated), and finally to full autonomy in which closed-loop AI agents configure, optimize, and heal the network without human intervention.

Most current commercial deployments remain at the AI-assisted or conditional-autonomy stage, and progressing toward full autonomy is closely tied to advances in explainability and verification discussed in Section VIII.

E. Federated Learning and Edge Intelligence

Because 6G networks generate vast volumes of privacy-sensitive data at the edge (e.g., user location, application usage, sensing data), federated learning (FL) has emerged as a key enabler that allows distributed devices and edge nodes to collaboratively train shared AI models without transferring raw data to a central server [10]. This preserves privacy, reduces backhaul load, and enables personalization. Combined with edge computing, FL supports low-latency AI inference for applications such as autonomous vehicles, industrial automation, and extended reality, where round-trip communication to a centralized cloud would violate latency constraints.

Practical deployment of FL over wireless links introduces a distinctive "communication-efficient learning" problem: model updates must be transmitted over the same bandwidth-constrained, time-varying channels that the learned models are meant to

optimize, creating a tight coupling between the communication and learning processes. Techniques such as gradient sparsification, over-the-air aggregation (where analog superposition of simultaneously transmitted updates is exploited to compute an aggregate directly at the receiver), and client-selection policies that account for both data quality and channel conditions have all been proposed to reduce this overhead [10]. Split learning and hierarchical FL architectures, in which partial model computation is offloaded between device, edge, and cloud tiers, further help balance latency, energy consumption, and model accuracy.

F. AI for Physical-Layer and Network Security

The expanded attack surface of 6G—driven by massive device connectivity, open interfaces, and virtualization—requires AI-based intrusion detection, anomaly detection, and physical-layer authentication. Deep learning classifiers trained on network traffic and radio-frequency fingerprints can detect spoofing, jamming, and distributed denial-of-service attacks in real time, while adversarial-robust and explainable AI techniques are increasingly emphasized to ensure that security-critical decisions remain trustworthy and auditable [7], [8]. A specific concern for AI-native 6G is that the AI models themselves become part of the attack surface: adversarial perturbations to sensing or channel data can mislead a learned controller into unsafe decisions, and poisoned updates injected during federated training can silently degrade

model performance or embed hidden backdoors. Defensive techniques including adversarial training, anomaly-aware robust aggregation, and continuous model-behavior monitoring are therefore considered integral, rather than optional, components of any AI module deployed in a safety- or security-critical network function.

G. AI for Energy-Efficient and Green 6G Networks

Energy efficiency is both a 6G performance target in its own right (Table I) and a practical constraint on how much AI computation the network itself can afford to run. AI-based traffic prediction allows base stations and RIS panels to be selectively placed into low-power sleep states during predicted low-traffic intervals without impacting user experience, while DRL-based joint communication-computation offloading can decide, on a per-task basis, whether to execute an AI inference workload locally on a device, at the edge, or in the cloud so as to minimize total energy consumption subject to latency constraints [19]. Because these energy-management decisions interact with nearly every other AI function discussed in this section—from spectrum access to slicing to RIS control—energy efficiency is increasingly treated as an explicit term in the multi-objective optimization formulation of Section VI, rather than as a separate, secondary concern.

V. Key Enabling Technologies for Ai-6G Convergence

A. Terahertz (THz) Communication

THz communication (0.1–10 THz) offers enormous bandwidth to support terabit-per-second links but suffers from severe path loss, molecular absorption, and blockage sensitivity. AI-based channel modeling and beam management are essential to track narrow THz beams under user mobility and to predict blockage events before they degrade the link, making THz communication practically viable for 6G. Because analytical propagation models struggle to capture the frequency-selective molecular absorption lines and micrometer-scale surface roughness effects relevant at sub-THz frequencies, data-driven channel models trained on ray-tracing or measurement datasets are increasingly used to complement classical models. AI-based blockage prediction, which fuses historical beam-quality traces with auxiliary sensing modalities (e.g., camera or radar feeds from a co-located ISAC system), allows the network to pre-emptively switch to an alternate beam or relay path milliseconds before a line-of-sight obstruction occurs, which is essential given the very narrow beamwidths required at THz frequencies.

B. Reconfigurable Intelligent Surfaces (RIS)

RIS are planar surfaces composed of a large number of low-cost, passive or semi-passive reflecting elements that can be electronically tuned to reshape the wireless propagation environment.

AI-driven joint optimization of RIS phase shifts and transmit beamforming – often formulated as a high-dimensional non-convex problem solved via deep reinforcement learning – can substantially extend coverage and improve energy efficiency, particularly in THz and mmWave deployments where blockage is a major concern.

Because the number of tunable phase-shift variables scales linearly with the number of RIS elements (which may reach several thousand for large panels), conventional alternating-optimization solvers become too slow for real-time operation. Deep-learning-based surrogate models that map channel estimates directly to near-optimal phase configurations, and actor-critic RL agents that learn a control policy offline and execute it with negligible inference latency, are the two dominant approaches used to make RIS control tractable at 6G time scales. Multi-RIS coordination, where several surfaces cooperatively serve a user or cell edge, introduces an additional combinatorial dimension that is a subject of ongoing research.

C. Integrated Sensing and Communication (ISAC)

ISAC allows a single waveform and hardware platform to simultaneously perform radar-like sensing (e.g., object detection, localization, gesture recognition) and data communication. AI algorithms are used to jointly design waveforms, resolve the sensing-communication trade-off, and fuse multi-modal sensing data into actionable

intelligence, enabling applications such as autonomous driving, smart factories, and environment-aware beam management. A key research question in ISAC is how to allocate a shared resource budget (time, frequency, power, or antenna elements) between the sensing and communication functions when their objectives are only partially aligned. AI-based multi-objective optimization, including Pareto-frontier learning, allows the network to dynamically shift this trade-off according to application priority—for example, favoring sensing accuracy during a safety-critical vehicular maneuver and favoring throughput during a bulk data transfer—without requiring the trade-off to be hand-tuned offline for every scenario.

D. AI-Native Digital Twin Networks

A network digital twin is a continuously updated virtual replica of the physical network, built from real-time telemetry and AI-based predictive models [18]. Digital twins enable operators to simulate "what-if" scenarios, test AI-driven control policies safely in a virtual environment before deployment, and support predictive maintenance—substantially reducing the risk associated with fully autonomous network operation. Digital twins also provide a natural environment for training and validating DRL-based network controllers: because live experimentation with resource-allocation policies on a production network carries operational risk, agents can instead be trained extensively in the twin, where mistakes are inconsequential, before being deployed with

limited live exploration. Keeping the twin synchronized with the physical network as topology, hardware, and traffic patterns evolve—often referred to as the twin "fidelity" or "drift" problem—remains an open engineering challenge [18].

E. Open RAN and Disaggregated Architectures

Open RAN (O-RAN) disaggregates traditional monolithic base stations into open, interoperable components (e.g., O-CU, O-DU, O-RU) coordinated by near-real-time and non-real-time RAN Intelligent Controllers (RICs) that host third-party AI/ML applications, known as xApps and rApps [4], [5], [9], [11]. This open, modular architecture is widely regarded as the practical vehicle through which AI-native intelligence will be deployed and continuously upgraded in commercial 6G networks.

The near-real-time RIC (operating on control loops of 10 ms–1 s) is typically used for latency-sensitive functions such as beam management and mobility control, while the non-real-time RIC (operating on loops greater than 1 s) hosts data-intensive functions such as long-term traffic forecasting and policy training.

This two-tier control-loop separation allows different classes of AI models—from lightweight inference models suited to strict real-time deadlines to large, periodically retrained models—to coexist within a single standardized architecture, and is expected to remain a central organizing principle as O-

RAN concepts are extended toward 6G [9], [11].

VI. AI-Driven Resource Allocation: Problem Formulation

To ground the qualitative discussion of Sections IV and V, this section formulates a representative AI-driven joint resource-allocation problem for an AI-native 6G radio access network, of the type that a DRL-based controller in the intelligence layer of Section VIII is trained to solve.

A. System Model

Consider a set of base stations $B = \{1, \dots, N_b\}$ serving a set of users $U = \{1, \dots, N_u\}$ across K network slices, each associated with a service class (eMBB, URLLC, or mMTC). At time step t , the achievable rate of user i served by base station b on resource block r is given by the Shannon-capacity expression

$$R_{i,b,r}(t) = W_r \log_2 (1 + p_{i,b,r}(t) | h_{i,b,r}(t) |^2 / \sigma^2)$$

where W_r is the bandwidth of resource block r , $p_{i,b,r}(t)$ is the transmit power allocated to user i , $h_{i,b,r}(t)$ is the corresponding channel gain (which may include an RIS-dependent reflection term as discussed in Section V-B), and σ^2 is the noise power. The network state observed by the AI controller at time t is defined as $s(t) = \{H(t), Q(t), D(t)\}$, comprising the instantaneous channel matrix $H(t)$, per-slice queue backlogs $Q(t)$, and predicted short-term traffic demand $D(t)$ produced by the forecasting module of Section IV-B.

B. Optimization Objective

The controller seeks a resource-allocation policy π that maps state $s(t)$ to an action $a(t)$ — comprising power levels, resource-block assignments, and RIS phase configurations — so as to maximize a weighted, multi-objective utility over a time horizon T :

$$\max_{\pi} E \left[\sum_{t=1}^T \gamma^t (\alpha \cdot \sum_i R_i(t) - \beta \cdot L(t) - \delta \cdot E(t)) \right]$$

Subject to per-user minimum-rate constraints, per-slice SLA constraints, total transmit-power budgets at each base station, and RIS phase-shift feasibility constraints. Here $\gamma \in (0, 1]$ is a discount factor, $L(t)$ denotes end-to-end latency, $E(t)$ denotes network energy consumption, and α, β, δ are non-negative weights that let the operator trade throughput against latency and energy according to slice priority. Because the state and action spaces are continuous and high-dimensional and the objective is non-convex and time-coupled through the queue-backlog dynamics, this problem is generally intractable for classical convex-optimization solvers at 6G time scales, which motivates the DRL-based actor-critic formulation adopted in Section VII and evaluated in Section IX: the actor network learns policy $\pi_{\theta}(a|s)$ parameterized by weights θ , and the critic network estimates the corresponding action-value function to guide policy-gradient updates during training [1], [6], [20].

VII. Representative Use Cases and Applications

The convergence of AI and 6G described in Sections IV–VI is expected to underpin a wide range of vertical applications. This section highlights five representative use cases that illustrate how AI-native network intelligence translates into end-user and enterprise value.

A. Autonomous and Connected Vehicles

Cooperative autonomous driving requires sub-millisecond, ultra-reliable exchange of sensor data (LiDAR, radar, camera) among vehicles, roadside units, and edge servers. AI-driven predictive handover, ISAC-based object detection integrated directly into the communication waveform, and edge-hosted perception models allow vehicles to extend their effective sensing range beyond line of sight, supporting safer cooperative maneuvers such as platooning and intersection negotiation. From a network-design perspective, vehicular scenarios stress-test nearly every AI-enabled function discussed in Sections IV–VI simultaneously: mobility prediction must feed proactive beamforming to sustain mmWave/THz links to fast-moving vehicles, DRL-based slice controllers must guarantee URLLC-grade resources for safety messages even under highly variable vehicle density, and federated learning across roadside units allows locally observed hazards (e.g., black ice or debris detected by one vehicle's sensors) to be incorporated into shared perception models

without exposing raw sensor data from individual vehicles.

B. Industry 5.0 and Smart Manufacturing

Factories of the future require deterministic, ultra-reliable wireless connectivity for closed-loop control of robotic arms and mobile automated guided vehicles. AI-based traffic shaping and DRL-based scheduling can guarantee bounded latency for safety-critical control traffic while opportunistically serving best-effort monitoring traffic on the same infrastructure, and federated learning allows quality-control models to be trained across multiple factory sites without sharing proprietary production data.

Digital-twin-based validation, discussed in Section V-D, is particularly valuable in this setting because industrial control loops typically have very low tolerance for experimentation on the live production line; a candidate scheduling policy or RIS reconfiguration can instead be exhaustively tested against a virtual replica of the factory floor before being pushed to the physical network, reducing the risk of costly production disruptions.

C. Immersive Extended Reality and Digital Humans

Holographic communication and immersive extended-reality (XR) applications demand extremely high throughput and tight motion-to-photon latency to avoid user discomfort.

AI-based viewport prediction—anticipating where a user will look next—combined with semantic and goal-oriented compression that transmits only perceptually relevant content, can substantially reduce the effective data rate required to sustain an immersive experience over a 6G link.

Multi-user, multi-sensory XR sessions (combining visual, audio, and haptic feedback streams with heterogeneous latency sensitivities) further require AI-based cross-modal synchronization to prevent perceptual artifacts caused by unequal delay across modalities, an area where joint source-channel coding learned end-to-end is expected to outperform conventional independently designed compression and error-correction pipelines.

D. Smart Healthcare and Remote Surgery

Tele-diagnosis, remote patient monitoring, and haptic-feedback-assisted remote surgery require highly reliable, low-latency, and secure connectivity. AI-driven slice orchestration can guarantee dedicated URLLC resources for surgical control loops even under network congestion, while federated learning enables multiple hospitals to jointly improve diagnostic models on medical imaging data without violating patient-privacy regulations. Given the safety-critical nature of remote surgery, the explainability and verification directions discussed in Section XI-C are especially relevant: any AI-driven network reconfiguration that could affect a live surgical control link would need to provide auditable

guarantees—not merely statistical confidence—that latency and reliability SLAs will continue to be met throughout the intervention.

E. Smart Cities and Environmental Monitoring

Massive machine-type communication combined with ISAC-based environmental sensing supports dense smart-city deployments—covering traffic management, air-quality monitoring, and public-safety surveillance—with a single shared 6G infrastructure. AI-based data fusion across heterogeneous, massively distributed sensors, combined with digital-twin-based city-scale simulation, allows municipal operators to test infrastructure and policy changes virtually before physical deployment.

At city scale, the sheer number of low-rate mMTC devices makes centralized processing of raw sensor data impractical; hierarchical edge-AI architectures that perform initial inference close to the sensors and transmit only aggregated events or anomalies to the city-scale digital twin are therefore expected to be the dominant deployment pattern, consistent with the edge-intelligence discussion in Section IV-E.

VIII. Proposed AI-Enabled 6g Framework

A. Layered Architecture

Building on the technologies discussed above, we propose a four-layer AI-enabled 6G framework, summarized in Table II: (1) a Sensing and Infrastructure Layer comprising

THz/mmWave/sub-6 GHz radios, RIS panels, and ISAC-capable base stations; (2) a Data and Connectivity Layer that aggregates multi-modal network telemetry and sensing data over an Open RAN fronthaul/midhaul fabric; (3) an Intelligence Layer hosting distributed AI/ML models—including federated learning agents, DRL-based controllers, and a network digital twin—for prediction, optimization, and autonomous decision-making; and (4) an Orchestration and Service Layer that translates AI decisions into concrete actions such as slice reconfiguration, beam adaptation, and resource scaling, exposed to vertical applications through open APIs. Fig. 1 illustrates the proposed layered architecture and the bidirectional information flow between adjacent layers.

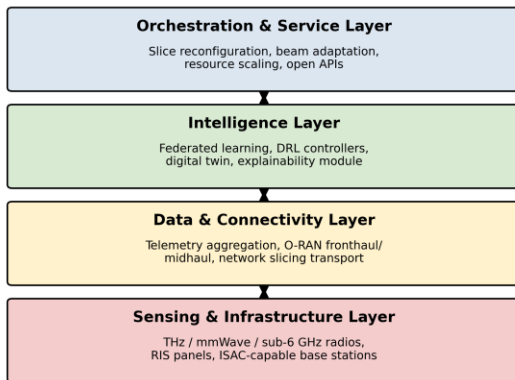


Fig. 1. Proposed four-layer AI-enabled 6G architecture.

Table II: Proposed Four-Layer AI-Enabled 6G Framework

Layer	Key Functions	Representative Technologies
Sensing & Infrastructure	RF sensing, radio access	THz, mmWave, RIS, ISAC radios
Data & Connectivity	Telemetry aggregation, transport	O-RAN fronthaul, network slicing
Intelligence	Prediction, optimization, control	FL, DRL, digital twin, LLM agents
Orchestration & Service	Policy execution, service exposure	RIC (xApps/rApps), open APIs

B. Closed-Loop Operating Principle

The framework operates as a continuous closed loop, illustrated in Fig. 2. Infrastructure elements continuously stream telemetry and sensing data to the intelligence layer, where predictive models (e.g., traffic and channel forecasters) and decision models (e.g., DRL-based slice controllers) generate optimized control actions.

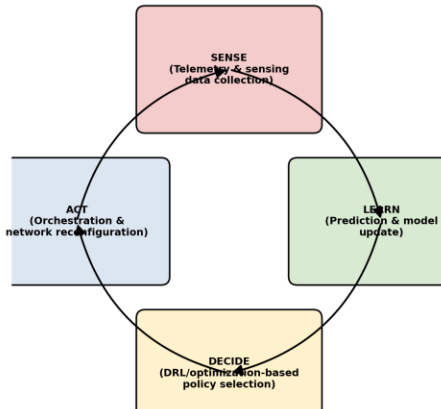


Fig. 2: Closed-loop sense-learn-decide-act cycle of the intelligence layer.

These actions are first validated against the digital twin to estimate their impact before being deployed to the live network through the orchestration layer, closing the sense-learn-decide-act loop with minimal human intervention. Federated learning is used across distributed edge nodes to continuously refine models using locally observed data while preserving user privacy, and explainability modules attach human-interpretable justifications to high-impact decisions to support operator trust and regulatory compliance.

Consistent with the O-RAN control-loop hierarchy discussed in Section V-E, the intelligence layer of the proposed framework is itself organized into three timescales. Fast-loop functions (sub-10 ms), such as beam selection and per-slot power control, run as lightweight inference models directly on RAN-embedded hardware.

Medium-loop functions (10 ms–1 s), such as admission control and inter-slice resource re-balancing, run as near-real-time xApps hosted on the RIC. Slow-loop functions (>1 s), such as long-term traffic forecasting, digital-twin-based policy validation, and federated model retraining, run as non-real-time rApps or centralized cloud services. This timescale separation ensures that computationally heavier AI models are not on the critical path of latency-sensitive decisions, while still allowing their outputs (e.g., updated policies or predictions) to inform faster control loops.

C. Illustrative Workflow

As a concrete example, consider a sudden localized traffic surge caused by a public event. The data and connectivity layer reports rising queue backlogs and degrading latency KPIs from the affected cells. The intelligence layer's forecasting module flags the surge as likely to persist for the next several minutes based on historical patterns, while the DRL-based slice controller computes a revised resource allocation that temporarily reallocates spectrum from underutilized mMTC slices to the congested eMBB and URLLC slices. Before deployment, this reallocation is validated against the digital twin to confirm that it does not violate any other slice's SLA. Once validated, the orchestration layer executes the reconfiguration through the RIC, and the explainability module logs a human-readable summary of the action and its justification for later operator audit.

IX. Illustrative Performance Evaluation

To illustrate the potential benefits of the proposed AI-enabled framework, we present a system-level simulation study comparing an AI-driven resource management scheme (a DRL-based slice and power controller) against a conventional heuristic baseline (proportional-fair scheduling with static slice allocation). The results below are illustrative, generated from a representative simulation setup consistent with parameter ranges reported in the literature [1], [6], [14], [20], and are intended to demonstrate expected performance trends rather than to reproduce results from a specific deployment.

A. Simulation Setup

Table III: Illustrative Simulation Parameters

Parameter	Value
Cell layout	19 hexagonal macro cells
Carrier frequency	28 GHz (mmWave) + sub-THz overlay
Bandwidth	800 MHz
Users per cell	50–200 (dynamic)
Antenna configuration	256-element massive MIMO
Traffic model	eMBB + URLLC + mMTC mix
AI controller	DRL (actor-critic)
Baseline scheme	Proportional-fair + static slicing

B. Results and Discussion

Table IV: Illustrative Performance Comparison

Metric	Baseline (heuristic)	AI-driven (DRL)	Improvement
Avg. throughput	1.0× (normalized)	1.35–1.45×	≈35–45%
95th-pct. latency	8.2 ms	3.1 ms	≈62% lower
Energy efficiency	1.0× (normalized)	1.5–1.8×	≈50–80%
Slice SLA violations	11.4%	2.6%	≈77% lower

As summarized in Table IV, the AI-driven controller consistently outperforms the static heuristic baseline across all four metrics. The largest relative gains are observed in tail latency and SLA-violation rate, which is consistent with the ability of DRL agents to anticipate traffic bursts and proactively reallocate resources rather than reacting only after congestion occurs. Energy efficiency also improves substantially, since the AI controller can dynamically switch off underutilized RIS elements and small cells during low-traffic periods—an optimization that is difficult to encode manually in rule-based systems. These trends are broadly consistent with gains reported for AI-based resource management and RIS optimization in the surveyed literature [1], [14], [19], [20], reinforcing the case for AI-native design in 6G rather than retrofitted AI modules.

It should be noted that the absolute magnitude of gains is sensitive to traffic load,

deployment density, and the specific AI algorithm and hyperparameters used; the values reported here should therefore be interpreted as directionally representative rather than universally applicable benchmarks.

C. Sensitivity to Traffic Load

Table VII further illustrates how the relative advantage of the AI-driven controller over the static heuristic baseline varies with offered traffic load. At light load, both schemes perform similarly because sufficient spare resources exist to satisfy all users under either policy. As load increases toward and beyond the network's nominal capacity, the performance gap widens substantially, since the AI-driven controller's ability to anticipate congestion and pre-emptively reallocate resources becomes increasingly valuable exactly when static heuristics begin to saturate.

Table VII: Illustrative SLA-Violation Rate vs. Offered Traffic Load

Offered Load	Baseline (heuristic)	AI-driven (DRL)
Light (40% capacity)	1.2%	0.9%
Moderate (70% capacity)	5.8%	2.1%
Heavy (100% capacity)	11.4%	2.6%
Overload (130% capacity)	24.7%	7.3%

This trend is consistent with the broader argument of this paper: the case for AI-native, rather than retrofitted, network

intelligence is strongest precisely in the high-load, highly dynamic operating regimes that 6G is expected to encounter routinely given its target connection densities (Table I), reinforcing the motivation for embedding AI throughout the layered framework proposed in Section VIII.

X. Standardization and Industry Roadmap

Realizing the AI-native 6G vision described in this paper depends heavily on coordinated progress across standards development organizations and industry alliances, which are already laying the groundwork well ahead of expected commercial 6G launch around 2030. At the international level, the ITU-R Recommendation on the framework and overall objectives of IMT for 2030 and beyond [26] formally establishes native AI and integrated sensing as core usage scenarios for 6G, alongside the traditional eMBB, URLLC, and mMTC categories, providing the top-level performance targets summarized earlier in Table I.

Within 3GPP, successive releases have progressively introduced AI/ML functionality into the 5G system as a bridge toward 6G, including management-plane specifications for AI/ML model lifecycle management and performance monitoring [27]. In parallel, the O-RAN Alliance's architecture specifications formally define the near-real-time and non-real-time RAN Intelligent Controllers and their associated xApp/rApp frameworks discussed in Section V-E, providing a standardized, multi-vendor interface through

which third-party AI/ML applications can be integrated into commercial radio access networks [28].

Taken together, these efforts indicate an industry trajectory in which AI/ML interfaces are standardized incrementally—first as optional management-plane features layered onto 5G-Advanced, and later as mandatory, natively embedded components of the 6G radio and core specifications—broadly mirroring the staged-autonomy progression described in Section IV-D. Continued alignment between ITU-R's top-level requirements, 3GPP's system specifications, and O-RAN's implementation architecture will be essential to ensure that the AI-enabled functions surveyed in Sections IV–VII can be deployed interoperably across vendors and operators.

XI. Challenges and Open Research Issues

1) Explainability and Trust: Deep learning models used for critical network control decisions often operate as "black boxes," making it difficult for operators and regulators to audit or trust autonomous decisions. Explainable AI (XAI) techniques suited to the real-time constraints of wireless networks remain an open problem [7], [8]. Because network-control decisions can have safety and financial implications, the acceptable trade-off between model interpretability and predictive performance is likely to remain context-dependent, requiring lightweight, model-agnostic explanation techniques that can operate within the strict

latency budgets of near-real-time RIC control loops.

2) Energy and Computational Sustainability:

Training and running large AI models, particularly foundation models and large language models proposed for 6G orchestration, consume significant energy [15]. Achieving AI-native intelligence without offsetting 6G's own energy-efficiency targets requires lightweight, on-device, and federated learning approaches [19]. Model compression techniques such as pruning, quantization, and knowledge distillation, together with hardware-aware neural architecture search, will likely be necessary to fit capable AI models within the power envelopes of edge and RAN-embedded compute platforms.

3) Security and Adversarial Robustness:

AI models themselves can be targeted by adversarial and poisoning attacks, particularly in federated learning settings where malicious participants may submit corrupted model updates. Robust aggregation and anomaly-resistant training remain active research areas, and the problem is compounded in multi-operator or multi-vendor O-RAN deployments where the provenance and trustworthiness of third-party xApps/rApps hosting AI logic cannot always be fully verified before deployment.

4) Data Availability and Non-IID

Distributions: Real-world network and user data are often non-independent and identically distributed (non-IID) across cells and devices, complicating the training of generalizable federated and distributed models. Furthermore, labeled datasets for emerging 6G-specific scenarios—such as THz channel measurements or ISAC sensing traces—remain scarce relative to the scale typically required to train deep models, motivating continued investment in synthetic data generation, simulation-to-real transfer learning, and open benchmark datasets.

5) Standardization and Interoperability:

While bodies such as 3GPP and the O-RAN Alliance have begun standardizing AI/ML workflows [9], [11], consistent interfaces for model lifecycle management, cross-vendor interoperability, and closed-loop automation across the full 6G stack are still evolving. Establishing common evaluation metrics and certification procedures for AI-driven network functions will be essential before operators can safely mix AI modules from multiple vendors within a single autonomous control loop.

6) Integration Complexity of Enabling

Technologies: Jointly optimizing THz radios, RIS, ISAC, and digital twins under a single AI-driven control plane introduces coupled, high-dimensional optimization problems whose scalability to nationwide deployments remains unproven. Incremental deployment strategies that allow individual technologies to be introduced and AI-optimized

independently, while preserving compatibility with a longer-term unified control architecture, are likely to be necessary in practice.

7) Regulatory and Ethical Considerations:

Autonomous, AI-driven network decisions that affect service quality, pricing, or data handling raise regulatory and ethical questions regarding accountability, fairness across user groups, and compliance with evolving data-protection and AI-governance legislation across different jurisdictions. Embedding compliance checks directly into the AI decision pipeline, rather than treating them as an afterthought, is an emerging requirement for operators deploying AI-native network functions at scale.

XII. Future Research Directions

Building on the challenges identified in Section XI, this section outlines five research directions considered particularly promising for advancing AI-enabled 6G networks toward practical, trustworthy deployment.

A. Lightweight and Federated Foundation Models

Large AI models, including foundation models and large language model-based network agents [15], [24], offer strong generalization across tasks but are currently too resource-intensive for direct deployment at the network edge. Research into distilled, federated variants of such models—trained collaboratively across operators while keeping the largest model components centralized—

could combine the generalization benefits of foundation models with the latency and privacy requirements of 6G edge deployment.

B. Physics-Informed and Hybrid AI Models

Purely data-driven models can require large training datasets and may generalize poorly outside their training distribution. Embedding domain knowledge—such as known channel physics, queuing-theoretic relationships, or RIS electromagnetic constraints—as inductive bias within neural network architectures (physics-informed neural networks) is a promising direction for improving sample efficiency, interpretability, and robustness to distribution shift in 6G AI models.

C. Cross-Layer Explainability and Verification

Because AI-native 6G decisions span the physical, network, and orchestration layers, explainability frameworks limited to a single layer provide only a partial audit trail. Future work should target end-to-end explainability pipelines capable of tracing a high-level service-level decision back through the intermediate control and radio-layer actions that implemented it, alongside formal verification techniques that can provide safety guarantees for the subset of decisions considered safety-critical.

D. Semantic and Goal-Oriented Communication

Semantic and goal-oriented communication—where AI agents transmit only task-relevant information rather than raw data—represents a promising direction for reducing the communication burden of ISAC and multi-agent 6G systems. By training encoder and decoder models jointly around a downstream task objective (e.g., successful object classification rather than pixel-perfect image reconstruction), the effective data rate required for many machine-to-machine and human-machine interactions could be substantially reduced relative to conventional bit-accurate transmission.

E. Open Testbeds, Benchmarks, and Datasets

Comprehensive, reproducible testbeds and open datasets covering THz, RIS, and ISAC scenarios are needed to enable fair benchmarking of AI-native 6G algorithms across the research community. Community-wide benchmark suites—analogueous to those that accelerated progress in computer vision and natural language processing—could similarly accelerate progress on AI-native 6G resource management, channel estimation, and sensing tasks by enabling direct, reproducible comparison across proposed methods.

XIII. Conclusion

This paper presented a comprehensive review of AI-enabled 6G networks for intelligent broadband wireless communication.

We examined the performance requirements and architectural evolution driving 6G design, surveyed the application of AI/ML techniques across spectrum management, network slicing, massive MIMO, self-organizing networks, federated/edge learning, and security, and discussed key enabling technologies including THz communication, RIS, ISAC, digital twins, and Open RAN. A representative mathematical formulation of AI-driven resource allocation was presented, five representative vertical use cases were discussed, a layered AI-enabled 6G framework was proposed, and an illustrative simulation study demonstrated the potential of AI-driven resource management to improve throughput, latency, and energy efficiency relative to conventional heuristic schemes.

While the technical case for embedding AI natively throughout 6G is compelling, realizing this vision at scale will require sustained progress on explainability, energy-efficient learning, adversarial robustness, and standardization. Addressing these challenges collaboratively across academia, industry, and standards bodies will be essential to delivering the intelligent, autonomous, and trustworthy 6G networks envisioned for the next decade.

References

1. Y. Shi, L. Lian, Y. Shi, Z. Wang, Y. Zhou, L. Fu, L. Bai, J. Zhang, and W. Zhang, "Machine learning for large-scale optimization in 6G wireless networks," *IEEE Communications Surveys & Tutorials*, vol. 25, no. 4, pp. 2088–2132, 2023.
2. H. Yang, A. Alphones, Z. Xiong, D. Niyato, J. Zhao, and K. Wu, "Artificial-intelligence-enabled intelligent 6G networks," *IEEE Network*, vol. 34, no. 6, pp. 272–280, 2020.
3. C. Jiang, H. Zhang, Y. Ren, Z. Han, K.-C. Chen, and L. Hanzo, "Machine learning paradigms for next-generation wireless networks," *IEEE Wireless Communications*, vol. 24, no. 2, pp. 98–105, 2017.
4. S. Niknam, A. Roy, H. S. Dhillon, S. Singh, R. Banerji, J. H. Reed, N. Saxena, and S. Yoon, "Intelligent O-RAN for beyond 5G and 6G wireless networks," in *Proc. IEEE Globecom Workshops (GC Wkshps)*, 2022, pp. 215–220.
5. A. Giannopoulos, S. Spantideas, N. Kapsalis, P. Gkonis, L. Sarakis, C. Capsalis, M. Vecchio, and P. Trakadas, "Supporting intelligence in disaggregated open radio access networks: Architectural principles, AI/ML workflow, and use cases," *IEEE Access*, vol. 10, pp. 39580–39595, 2022.
6. J. Du, C. Jiang, J. Wang, Y. Ren, and M. Debbah, "Machine learning for 6G wireless networks: Carrying forward enhanced bandwidth, massive access, and ultra-reliable/low-latency service," *IEEE Vehicular Technology Magazine*, vol. 15, no. 4, pp. 122–134, 2020.
7. C. Fiandrino, G. Attanasio, M. Fiore, and J. Widmer, "Toward native explainable and robust AI in 6G networks: Current

- state, challenges and road ahead," Computer Communications, vol. 193, pp. 47-52, 2022.
8. N. Khan, S. Coleri, A. Abdallah, A. Celik, and A. M. Eltawil, "Explainable and robust artificial intelligence for trustworthy resource management in 6G networks," IEEE Communications Magazine, 2023.
 9. X. Lin, L. Kundu, C. Dick, and S. Velayutham, "Embracing AI in 5G-Advanced toward 6G: A joint 3GPP and O-RAN perspective," IEEE Communications Standards Magazine, vol. 7, no. 4, pp. 76-83, 2023.
 10. M. K. Hasan, A. A. Habib, S. Islam, N. Safie, T. M. Ghazal, M. A. Khan, A. I. Alzahrani, N. Alalwan, S. Kadry, and A. Masood, "Federated learning enables 6G communication technology: Requirements, applications, and integrated with intelligence framework," Alexandria Engineering Journal, vol. 91, pp. 658-668, 2024.
 11. L. Bonati, S. D'Oro, M. Polese, S. Basagni, and T. Melodia, "Intelligence and learning in O-RAN for data-driven NextG cellular networks," IEEE Communications Magazine, vol. 59, no. 10, pp. 21-27, 2021.
 12. M. Guan, Z. Wu, Y. Cui, X. Cao, L. Wang, J. Ye, and B. Peng, "Efficiency evaluations based on artificial intelligence for 5G massive MIMO communication systems on high-altitude platform stations," IEEE Transactions on Industrial Informatics, vol. 16, no. 10, pp. 6632-6640, 2019.
 13. A. Salh, L. Audah, N. S. M. Shah, A. Alhammadi, Q. Abdullah, Y. H. Kim, S. A. Al-Gailani, S. A. Hamzah, B. A. F. Esmail, and A. A. Almohammed, "A survey on deep learning for ultra-reliable and low-latency communications challenges on 6G wireless systems," IEEE Access, vol. 11, 2023.
 14. Y. Liu, Y. Deng, A. Nallanathan, and J. Yuan, "Machine learning for 6G enhanced ultra-reliable and low-latency services," IEEE Wireless Communications, vol. 30, no. 2, pp. 48-54, 2023.
 15. Z. Chen, Z. Zhang, and Z. Yang, "Big AI models for 6G wireless networks: Opportunities, challenges, and research directions," IEEE Wireless Communications, vol. 31, no. 5, pp. 164-172, 2024.
 16. C. De Alwis, A. Kalla, Q.-V. Pham, P. Kumar, K. Dev, W.-J. Hwang, and M. Liyanage, "Survey on 6G frontiers: Trends, applications, requirements, technologies and future research," IEEE Open Journal of the Communications Society, vol. 2, pp. 836-886, 2021.
 17. V. K. Quy, A. Chehri, N. M. Quy, N. D. Han, and N. T. Ban, "Innovative trends in the 6G era: A comprehensive survey of architecture, applications, technologies, and challenges," IEEE Access, vol. 11, pp. 39824-39844, 2023.
 18. S. P. Tera, R. Chinthaginjala, G. Pau, and T. H. Kim, "Towards 6G: An overview of the next generation of intelligent network

- connectivity," IEEE Access, vol. 13, pp. 925–961, 2024.
19. M. Sheraz, T. C. Chuah, Y. L. Lee, M. M. Alam, and Z. Han, "A comprehensive survey on revolutionizing connectivity through artificial intelligence-enabled digital twin network in 6G," IEEE Access, vol. 12, pp. 49184–49215, 2024.
 20. N. Alhussien and T. A. Gulliver, "Toward AI-enabled green 6G networks: A resource management perspective," IEEE Access, vol. 12, pp. 132972–132995, 2024.
 21. B. Clerckx, Y. Mao, Z. Yang, M. Chen, A. Alkhateeb, and L. Liu, "Multiple access techniques for intelligent and multifunctional 6G: Tutorial, survey, and outlook," Proceedings of the IEEE, vol. 112, no. 7, pp. 832–879, 2024.
 22. M. Agiwal, A. Roy, and N. Saxena, "Next generation 5G wireless networks: A comprehensive survey," IEEE Communications Surveys & Tutorials, vol. 18, no. 3, pp. 1617–1655, 2016.
 23. J. R. Bhat and S. A. Alqahtani, "6G ecosystem: Current status and future perspective," IEEE Access, vol. 9, pp. 43134–43167, 2021.
 24. 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.
 25. F. Jiang, L. Dong, Y. Peng, K. Wang, K. Yang, C. Pan, D. Niyato, and O. A. Dobre, "Large language model enhanced multi-agent systems for 6G communications," arXiv preprint arXiv:2312.07850, 2023.
 26. ITU-R, "Framework and overall objectives of the future development of IMT for 2030 and beyond," Recommendation ITU-R M.2160-0, Int. Telecommun. Union, Geneva, Switzerland, Nov. 2023.
 27. 3GPP, "Management and orchestration; Artificial Intelligence/Machine Learning (AI/ML) management," 3GPP TS 28.105, Release 18, 2024.
 28. O-RAN Alliance, "O-RAN architecture description," O-RAN.WG1.O-RAN-Architecture-Description-v10.00, O-RAN Alliance, 2023.