AI Native 6G System Architecture

How the AI domain/AI plane is shaping the new 6G core network

Authors: Gabriel Brown and Ruth Brown

August 2026

Service intent and AI native 6G

3GPP has already made three attributes of the system architecture clear in the 6G Study Phase (Release 20):

- 6G will be standalone from Day 1.
- 6G will be AI native.
- 6G will be service-aware.

In September 2026, 3GPP will make critical decisions on 6G RAN, core, and end-to-end system architecture at its plenary meetings. Ahead of that meeting, this article discusses how the three high level concepts listed above might converge to shape the 6G core architecture to be defined formally in Release 21 (running from 2027 through 2029). The article also focuses on key issues (KIs) being addressed in 3GPP SA2 related to making AI a native part of the core architecture. It highlights proposed solutions to the KIs from leading industry players.

To set the context behind the three high level themes:

- **Standalone Day 1 requires a discrete 6G core network (CN).** What that means in practice is an open question. Is it a lightly upgraded version of a 5G core, a heavily evolved 5G core with lots of new 6G functions, or an entirely new core with new infrastructure, functions, protocols, and interfaces?
- **6G will be an AI native system**. A new AI domain will be integrated into the CN in all the above scenarios. This is sometimes also described as an “AI plane” that will run across the system architecture to augment the classic control and user planes. The details of this AI domain/AI plane are a work in progress.
- **Service-aware networks**. This is more nuanced because so-called intelligent networks have historically been outpaced by innovation in devices and applications. Nevertheless, new agentic and token-driven AI traffic profiles are likely to make service awareness important in 6G. (For background, see this 3GPP article on “Service characteristics awareness in the RAN,” July 2026).

Service intent is the concept that draws these three threads together. The idea behind service intent is that the device signals its connectivity requirements to the network (for a given application), and the AI plane acts as an orchestration engine to configure the network accordingly. Crucially, service intent is driven by the end-user application—that is, what the customer wants.

In this architecture, the AI orchestrator controls fleets of agents both operating in the new AI domain and embedded within network functions (NFs; see **Figure 1** below). Service intent is thus central to the 6G core and AI plane/AI domain discussions.

Mobile AI traffic profiles

One uncertainty for 6G designers is the impact mobile AI services will have on cellular networks. AI is changing computing and networking *in general*.

The impact on cellular networks from AI is, *so far*, relatively benign—at most just a few percent of traffic (typically text-based). However, AI is an era-defining technology, and 6G must anticipate and plan for it.

At some stage, an explosion in the number of AI-enabled mobile endpoints (phones, cameras, cars, robots, glasses, wearables, sensors, etc.) will inevitably affect the profile of mobile network traffic.

Plausibly, 6G network designers should expect AI to

- Generate greater data volumes—especially in the uplink due to machine vision applications (self-driving vehicles are already showing this).
- Enable agent-based communications that do not mirror human usage patterns in terms of busy hour, session duration, concurrent connections, and so on.
- Create demand for edge inference (network compute) for power-constrained devices, such as XR glasses, robots, and drones, and for 6G applications, such as sensing.

What this all means for traffic is hard to anticipate and to design for. Network designers must consider (for example) the number and length of sessions, the size of “token blocks,” latency requirements, the location of compute/inference, and the implications for “east-west” traffic within the network.

Core architecture evolution

The ongoing 3GPP 6G architectural framework discussions are based on Release 20 technical studies, such as TR.22.870 (*SA1 Service Requirements*) and TR.23.801 (*Study on Architecture for 6G Systems; Stage 2*). These studies evaluate 6G architecture proposals submitted by the ecosystem and identify KIs to address ahead of work on formal standards development in Release 21.

There is broad agreement that 6G will be an AI native system. AI will be embedded into all elements and layers of the network (rather than deployed as an overlay functionality as in 5G), and much of this will be defined in the CN.

The 6G AI native system architecture will be formed around the following principles:

- **Service-based-architecture (SBA):** 5G core SBA framework is the starting point for discussion and appears likely to be reused for 6G core.
- **Standalone 6G architecture**: Control and user plane of the 6G RAN connects to a single CN type (i.e., 6G CN) on Day 1.
- **Preserve access-core split:** Functional separation of the RAN and CN to ensure multi-vendor interoperable interfaces and roaming.
- **Duplication:** Avoid function duplication between 6G RAN and CN. Minimize interdependencies between NFs when introducing a new feature or procedure for an NF.
- **Intelligent network-to-user equipment (UE) collaboration:** UE behavior prediction and network collaboration to enable service resilience, optimization, and adaptive provisioning.
- **Network AI agent:** Intent-based service provisioning, orchestration, and dynamic optimization of 3GPP services, as well as interaction with UE/user/third-party agent and dynamic optimization of 3GPP services.
- **IP Multimedia Subsystem (IMS)-related capabilities:** Intelligent Communication Assistant services to enhance user experience and service delivery.

5G interworking for 6G Day 1

Given the imminent 3GPP meetings in September, it would be rash to make predictions on what the 6G core and system architecture will look like. Nevertheless, an approximate view is that for most early deployments (say, in 2029–31), the 6G core will be based on an evolved 5G core infrastructure enhanced with some new 6G core control-plane functions and a new AI domain.

For example, a new “6G AMF” may be introduced alongside combo 5G/6G nodes such as SMF, UPF, etc. (or some variation of new and combo nodes). Similarly, because voice and emergency calling are Day 1 requirements, this means reusing the 4G/5G IMS core. The initial 6G SA core will interwork with the existing 5G core, then evolve and refresh independently over time.

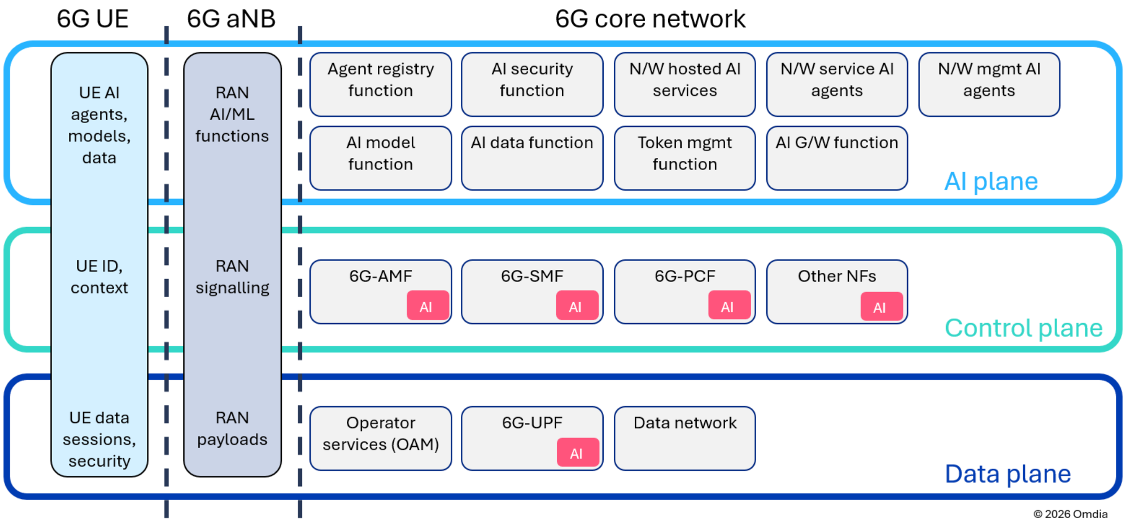
To a degree, this approach risks creating a two-phase 6G akin to the non-standalone (NSA) to standalone (SA) transition observed in 5G. However, the difference is that the architecture is SA from Day 1 and therefore the core can evolve quasi-independently of the RAN.

The 6G core AI domain/AI plane

The 6G architecture is likely to formalize a new AI domain in the CN. Based on Omdia’s reading of 3GPP submissions and other market commentary, a draft architecture appears to be emerging, as shown in **Figure 1**. In this architecture, NFs in the control and user planes are built with embedded agents that are orchestrated by the new agentic AI plane.

The emerging 6G AI domain closely mirrors existing hyperscaler (e.g., Google Gemini Enterprise Agent Platform, AWS Bedrock AgentCore, Microsoft Azure AI Foundry) and telecom vendor agent fabrics (e.g., Amdocs Agentic Operating System, Netcracker Agentic AI Platform).

Figure 1: Emerging AI architecture for 6G core



Source: Omdia

Within the telecom domain, this architecture is aligned with the TM Forum's new Open Digital Architecture (ODA) AI-native canvas. The TM Forum's AI-native canvas provides the following functionality:

- **AI models access:** Foundational generative AI (GenAI), large language models (LLMs), embedded models, multi-model models, etc.
- **AI gateway:** Provides auditing, automated guardrails, and cost management
- **Agent registry:** Detailing all agents, tools, and high risk use cases
- **Data:** Collection, management, knowledge bases, and distribution
- **Security management:** Agentic interaction security to evaluate session level and user requests, immutable log, etc.

In practical terms, this architecture builds on the agentic AI for IT operations (AIOps) model emerging organically in 5G networks today. An agentic network management layer is already being built by leading operators. Some vendors are evolving toward common data management and actuation for core and RAN.

The case for an all-new AI native 6G core

There is a strong technical case for an entirely new 6G core built as AI native from the ground up for Day 1 deployment. For example, the disaggregated architecture could align more closely to state-of-the-art AI compute, and NFs could be defined as larger functional blocks and connected over a native agent fabric. Automation, monitoring, assurance, and so on would follow suit. Potentially, an all-new AI native core would allow the network to match service intent more efficiently to network resources and deliver optimal customer outcomes.

However, 2030 is too soon operationally and financially for operators to make such a large and radical shift in core infrastructure and operations. It is essential to keep in mind how critical the core is to network availability and resiliency—the reliability of cloud native 5G SA core was hard-won. No serious operator will risk compromising service availability.

Ironically, an issue with the evolution argument is that 5G core (and the service-based architecture) is now often viewed as overly complex, too brittle/fragile, and overly disaggregated. Nonetheless, it is the baseline the industry must work from.

It is also too early in the cycle to know how AI technology will affect mobile networks. Right now, for all the hype, the service and revenue impacts from AI are still small. For operators, there is some logic in not committing to a new core architecture too soon.

An all-new 6G core is thus appealing in theory, but is unlikely to be widely implementable in practice in the 2030 timeframe in a manner that is scalable and deployable for the global mainstream. Breakthrough innovations, or a change in industry sentiment, may yet prove otherwise, and the authors of this paper would be delighted to be wrong about this prediction.

Some argue that retrofitting AI to an evolved 5G architecture and calling it 6G SA is a losing proposition from the start. To use an automotive analogy, evolving an existing 5G core for AI native 6G is like building an electric vehicle (EV) on a combustion engine chassis. It works, but the EV native design will ultimately win. To extend the analogy further, one can only build a successful EV platform if one also builds the production line to produce it.

This automotive analogy does not map perfectly to 6G because nearly all mobile operators run brownfield infrastructure. Yet, it shows the need for vendors and operators to balance ambitious innovation with practical implementation constraints when developing AI native 6G core technology.

Key issues for AI integration

The 3GPP SA2 working group is actively addressing AI integration through its study on architecture for 6G systems TR 23.801. This work establishes that AI for the 6G architecture shall be multi-vendor interoperable, reliable, and sustainable. TR 23.801 identifies four KIs that specifically address AI integration with the CN. Major vendors and operators are proposing solutions that reflect their technical approaches to AI native design.

The four KIs that relate most directly to AI for the 6G architecture are:

- **AI for 6G architecture (KI #18)** studies AI agents and frameworks. Critical study areas include handling intent, non-intent-based requests, dynamic procedure composition, AI-capable NFs, network AI capabilities access (internal and external), performance monitoring, operator control and autonomy, roaming, and 5G-6G AI interoperability.
- **6G Network for AI (KI #19)** addresses AI agent interactions and AI service provisioning. There are five study items: UE AI agent discovery on different UEs via 6G network(s), UE AI agent communication (on different UEs via 6G networks), network capability exposure enhancement, 6G CN AI service provisioning, and system impacts from AI traffic characteristics.
- **Data Framework (KI #21)** will ensure compatibility for data handling. The main study items include high level use case identification (e.g., AI/machine learning [ML] in the core, UE data collection for UE-sided model training, sensing applications), data framework functionalities (discovery, collection, labeling, storage, exposure, etc.), and system impacts (user consent, privacy, security, and governance).
- **6G Computing Support (KI #22)** across UE, CN, and application servers. Key aspects include architecture requirements and resource management, authorization mechanisms for UEs and AFs, coordination of communication and computing, service continuity and quality of service (QoS) considerations, and compute site discovery and selection/reselection.

Core network AI proposals

The initial 6G workshop at the 3GPP plenary meeting in April 2025 collected high level visions, which fed directly into the study item description. Subsequent working group meetings across the SA2 core system architecture study have begun more granular technical work, solution drafting, and proposal debate.

Submitted change requests (CRs) for the TR.23.801 working draft address the KIs identified above. **Table 1** highlights a non-exhaustive selection of solutions to the KIs proposed by key vendors and operators. The proposals span AI-embedded NFs, separate AI domains/environments, agent-centric orchestration, and much more.

Table 1: Comparison of emerging AI architectural proposals

	Design philosophies	Intent handling	Description
Skill/capability-centric			
Google S2-2602000	Skills-based agentic architecture	Capability discovery to fulfill service intent	• Agentic service stratum (i.e., proposing new structural layer, dual-bus architecture) • NF skills register (no API needed to invoke capability) • New NFs, including agentic capability repository function, agentic service function, skill registration, and management functionality • Update to NFs to make them skill aware (e.g., 6G-AMF, 6G-SMF, 6G-PCF)
AI-embedded NFs			
Samsung, AT&T S2-2601951	Distributed AI-embedded NF capabilities	AI-embedded NFs	• Distributed AI capabilities with embedded AI native NFs within the 6G core to work between AI and non-AI entities • AI-enabled NFs able to exchange inference outputs, model parameters, and context via SBI • NFs (e.g., PCF, SMF, UPF, etc.) become AI native; RAN node and OAM can work in conjunction to train joint AI model and perform joint inference
New AI domain/environment			
Nokia S2-2601946	Agentic overlay architecture	Agentic domain	• Separate, independent NF and agentic domains • Tool translator between agents and core functions • Deterministic NFs; may optionally embed AI agent functionality
Ericsson, T-Mobile USA, Deutsche Telekom S2-2602487	Separate (and optional) AI domain	AI domain	• Optional 6G AI domain • AI control function orchestrating intent; operator control on whether AI is used for intent • Standardizing agentic services interface via existing SBI • Evolved 5G NFs (AI-enabled)

	Design philosophies	Intent handling	Description
Qualcomm S2-2601951	Dual domain agentic architecture	Network agent + skills	• Strict separation between AI agent and 3GPP procedures and service domain • N/W AI agents expose agentic skills and capabilities • Within agent domain: AI agent proxy, subscription database, N/W AI agents, tool adapter
NVIDIA S2-2602248	AI plane architecture	AI domain	• AI plane logical architecture layer • Introduces new NFs: AI-GWF, AARF, AISF, network service AI agents, network-hosted AI services, network management, AI agents • 6G CN NFs may optionally support agentic operations via AI Control Interface (ACI)
Agent-centric orchestration			
China Telecom S2-2602211 S2-2602212	Hierarchical multi-agent AI architecture	Hierarchical intent decomposition	• Embedded AI inference in 6G NFs • AI control N/W function, service decomposition agent, specialized agents (domain-specific) • Hierarchical AI agents, with multi-level intent decomposition
Inter Digital S2- 2601914	Tool-enhanced agent-orchestrated architecture	Agent reasoning + tool orchestration	• New AI agent NF on existing SBI • Also introduces a network tool function that provides AI tools and capabilities, and can be accessed by UE or AF through the NF AI Agent (autonomous reasoning, tool discovery, and orchestration)
Huawei & China Mobile S2-260185	AI agent workflow architecture	Agent workflow orchestration	• Introduce agentic SBI (supports A2A and MCP interactions) • Embeds tools into existing NFs (tool-based design) • Agents include: system agent (intent), data AI agent, agent registration function

Source: Omdia

Conclusion: End users are the North Star

The potential for AI services is immense. This technology will expand the addressable market for mobile wireless technology and network operators and create value for end customers. Wireless networks are fundamental to delivering services to AI-enabled endpoints and to realizing the full utility of AI in the wider economy and society.

The critical questions are about when and how this happens. Service intent (what users/customers want) is a useful guiding principle to determine what the architecture should look like, how to build it, and how to operate 6G networks.

Based on the current industry direction, early 6G core implementations (2029–31) will be based on an evolved, cloud native 5G core infrastructure enhanced with new 6G NFs and a discrete AI domain/plane.

The emerging AI domain aligns with established hyperscaler and telecom vendor agent platforms and appears set to be specified in Release 21. In terms of the wider 6G core, the mobile network technology ecosystem is focused on the September 2026 3GPP meetings to see whether the industry commits to a more evolutionary or revolutionary 6G architecture.

Appendix

References

- *Study on 6G Use Cases and Service Requirements* (TR.22.870), 3GPP (ongoing)
- *Study on Architecture for 6G System; Stage 2* (TR.23.801), 3GPP (ongoing)
- SA2 Working Group, 174, Malta (April 2026) change requests, 3GPP (ongoing)
- *6G Progress Report*, GSMA (May 2026)
- TM Forum AI-Native Open Digital Architecture (ODA) Canvas (August 2026)

Omdia (formerly Heavy Reading) and Light Reading publications

- Gabriel Brown, *Set the Right Migration Path: Defining a 6G Architecture* (September 2025)
- Gabriel Brown, *6G will be standalone (SA) on day 1* (September 2025)
- Gabriel Brown, *AI Native 6G and Context-Aware Communications* (May 2026)
- Ruth Brown, *Analyst insight: 6G coming into focus* (June 2026)
- Ruth Brown, *The architecture shift reshaping the telco industry* (July 2026)

Gabriel Brown, Senior Principal Analyst, Mobile Networks, GTM Telecom Insights and Advisory

Ruth Brown, Senior Principal Analyst, Mobile Networks, GTM Telecom Insights and Advisory

Omdia consulting

Omdia is a market-leading data, research, and consulting business focused on helping digital service providers, technology companies, and enterprise decision makers thrive in the connected digital economy. Through our global base of analysts, we offer expert analysis and strategic insight across the IT, telecoms, and media industries.

We create business advantage for our customers by providing actionable insight to support business planning, product development, and go-to-market initiatives.

Our unique combination of authoritative data, market analysis, and vertical industry expertise is designed to empower decision-making, helping our clients profit from new technologies and capitalize on evolving business models.

Omdia is part of Informa TechTarget, a B2B Materials information services business serving the technology, media, and telecoms sector. The Informa group is listed on the London Stock Exchange.

We hope that this analysis will help you make informed and imaginative business decisions. If you have further requirements, Omdia's consulting team may be able to help your company identify future trends and opportunities.

Get in touch

www.omdia.com
askananalyst@omdia.com