

From AI innovation to AI at scale with Ericsson Telco IT AI Engine



**NO NONSENSE
OSS/BSS.**

Sell. Deliver. Get paid.



DC 62 143



150 years

Executive summary

Pressure to transform end-to-end telecom operations across [operations and business support systems \(OSS/BSS\)](#), IT, and networks is intensifying. Our challenge as an industry is to deliver real-time, digital-first customer experiences while enabling autonomous, data-driven networks. This urgency is being driven by rapid advances in technologies such as 5G, cloud-native core, Open RAN, and edge computing, along with unprecedented growth in data volumes from connected devices and digital services. At the same time, we must manage the increasing complexity of multi-vendor, software-defined, and highly distributed environments. In this context, operationalizing artificial intelligence (AI) at scale is no longer optional. It is critical to enable real-time decision-making, predictive operations, closed-loop automation, and self-healing networks.

As network complexity increases and service expectations continue to rise, AI-driven automation has become a top strategic priority across telco domains, with predictive AI, generative AI (GenAI), and agentic AI are expected to become deeply embedded in next-generation telco architectures. However, despite strong industry momentum, many AI initiatives remain confined to proof-of-concept stages, running on siloed, one-off stacks that are difficult to integrate, scale, or trust in live operations. This is mainly due to fragmented data landscapes, disconnected AI platforms across domains, integration and interoperability barriers, limited operational AI skills, and growing concerns around trust, explainability, security, and governance.

To move beyond fragmented pilots and realize AI's full operational potential, we need a unified, scalable, and trusted

AI platform purpose-built for telco environments. Designed to address these needs, [Ericsson Telco IT AI Engine](#) provides a cloud-agnostic, enterprise-grade AI platform spanning machine learning operations (MLOps), large language model operations (LLMOps), and agent operations (AgentOps). It unifies the design, deployment, and operation of machine learning (ML), generative AI, and agentic AI applications across OSS/BSS, IT, and network domains. Telco IT AI Engine standardizes the end-to-end AI lifecycle and integrates seamlessly with data platforms and enterprise systems. Telco IT AI Engine is a modular AI platform that allows customers and Ericsson OSS/BSS teams to accelerate AI innovation, reduce time to market for AI solutions, and establish the foundation for AI-first, autonomous operations across telecom domains.



Transform telco operations through AI-driven automation

Modernization of networks and operations to support 5G, IoT, edge-enabled services, B2B/B2B2X business models and network slicing is well underway, while meeting rising expectations for always-on, digital-first experiences and assured service levels. To deliver these, we must bring together cloud-native network and IT platforms with data and AI into cohesive, end-to-end architectures spanning OSS/BSS, core, radio access network (RAN), and transport.

AI has the potential to transform telecom customer service from reactive, generic, and manual to proactive, personalized, and highly automated. This can significantly improve both customer experience and operational efficiency. Recent industry research from [Omdia](#) shows that automating operations with AI is the top OSS/BSS priority for CSPs, as shown in Figure 1. While this research is focused on OSS/BSS, the same AI-driven automation priorities extend into network operations and broader telco transformation. AI is no longer a niche add-on. It is becoming a core design principle for next-generation telco architectures. It's time to embed predictive AI, GenAI, and agentic AI across multiple domains, including:

- **Network operations and assurance:** Predictive maintenance, anomaly detection, root cause analysis, intent-based networking, and closed-loop automation.
- **Customer and business operations:** Customer care, billing, charging, product catalog, order management, and partner ecosystem management.
- **Service lifecycle and experience management:** Quality of experience, service level agreement management, service assurance, and service optimization.

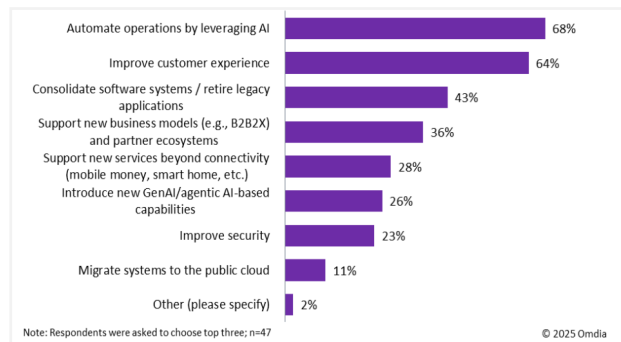


Figure 1: CSPs' top OSS/BSS priorities

Across these domains, AI-driven automation and operations are becoming central to transformation agendas. High-impact use cases such as root cause analysis, predictive maintenance, closed-loop assurance, intent-based operations, and AI-driven customer engagement can improve mean time to recover, reliability, net promoter score, and operational efficiency. GenAI and agentic AI are expected to evolve from productivity and decision-support tools into key enablers of autonomous, cloud-native networks and services.

However, several barriers continue to slow the journey from proof of concept to production and scale:

- **Data silos and limited data intelligence:** Fragmented data across OSS/BSS, IT, and network domains limits the ability to feed trusted, real-time data into AI systems.
- **Shortage of operational AI and LLM skills:** Limited expertise makes it difficult to industrialize AI, apply best practices, and govern models consistently across teams and domains.
- **Organizational silos and fragmented platforms:** Isolated ownership of AI initiatives and duplicated tooling often lead to bespoke pilots that are

difficult to reuse or scale across the organization.

- **Integration and interoperability challenges:** AI components are often tightly coupled to specific tools or environments, making it difficult to integrate them into heterogeneous, multi-vendor architectures and existing operational workflows.
- **Lack of standard frameworks and reusable patterns:** Inconsistent approaches to MLOps, LLMOps, and multi-agent systems result in duplicated engineering, variable quality, and high operational overhead for each new use case.
- **Limited trust in AI outcomes:** Without robust observability, explainability, safety guardrails, and data-protection mechanisms, teams remain hesitant to allow AI to influence or automate critical operations.

To move forward, we need unified data and AI platforms built on reusable patterns, supported by partners with deep telecom expertise. This combination is essential to design, industrialize, and operationalize AI across the full telco stack, from OSS/BSS and IT to network operations and service assurance.

Ericsson Telco IT AI Engine: Build a robust foundation for AI-driven automation and operations

With priorities clear and barriers persistent, we need a production-grade foundation that connects data, AI, and action across OSS/BSS, IT, and network domains. Ericsson Telco IT AI Engine provides that foundation, moving AI from isolated pilots to trusted, scalable operations. It brings together standardized MLOps, LLMOps, and AgentOps capabilities—eliminating silos, reducing duplication, and enabling consistent, scalable AI deployment across domains.

Ericsson Telco IT AI Engine is a cloud-agnostic AI platform designed for rapid development and deployment of ML, GenAI, and agentic AI applications across Ericsson telco domains. It serves as a unified intelligence layer that standardizes and orchestrates the complete AI lifecycle: from data preparation and model training through inference, monitoring, and autonomous operations.

Ericsson Telco IT AI Engine is built on a robust, multi-layered architecture delivering comprehensive, production-ready capabilities:

- **Enterprise-grade MLOps, LLMOps, and AgentOps platform:** Supports end-to-end development, deployment, monitoring, and lifecycle management of ML, GenAI, and multi-agent AI applications, all within a single operational framework.
- **Multi-layered architecture:** Comprises of a customer-facing user interface, intelligent orchestration layer, flexible storage and data layer, and seamless integration layer, enabling deployment flexibility and domain-specific customization without sacrificing operational coherence.

- **Seamless OSS/BSS and network integration:** Telco IT AI Engine seamlessly integrates with OSS/BSS and network platforms, acting as an intelligence layer that enhances existing operations. It supports flexible data integration, combining historical batch processing for model training with near-real-time and streaming pipelines

for event-driven insights. The platform exposes standardized model-serving APIs for predictions, anomaly detection, and decision support, and enables embedding GenAI assistants and agents directly into workflows to drive contextual insights and automation at the point of action.

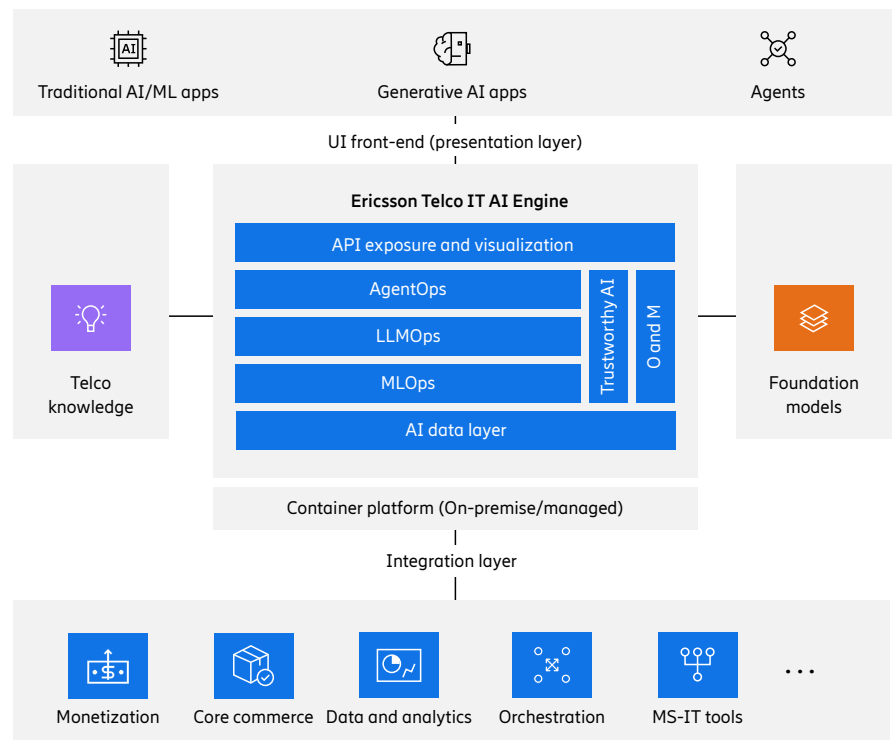


Figure 2: Ericsson Telco IT AI Engine functional architecture

Ericsson Telco IT AI Engine core capabilities

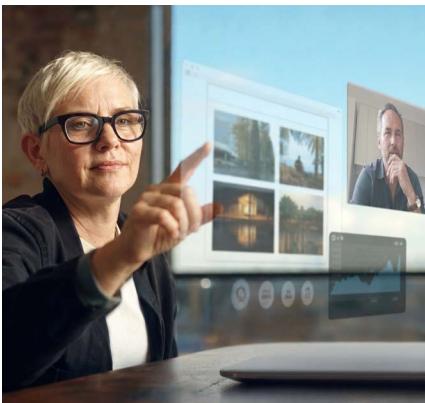


MLOps capabilities: Telco IT AI Engine provides an end-to-end MLOps stack to build, deploy, and manage ML models across OSS/BSS, network, and IT domains, supporting both traditional and advanced analytics use cases:

- End to end ML pipeline orchestration with distributed data exploration, preprocessing, model training, evaluation, deployment and lifecycle management
- Configuration-driven pipelines for common ML scenarios using prebuilt

modules (for example anomaly detection, forecasting, correlation analysis) to accelerate development for non specialist teams

- Scalable compute with support for distributed processing and training (for example Spark based runtimes) to handle large scale telco datasets such as network counters, logs, telemetry, customer, and IT data



LLMOps and generative AI capabilities:

Telco IT AI Engine offers a comprehensive environment to develop and operationalize generative AI applications:

- Prebuilt base images, libraries, and software development kits (SDK) to accelerate development and reduce time to value
- A scalable vector database engine to store knowledge bases and support retrieval-augmented generation (RAG) pipelines for documentation, topology, configuration, and telemetry, enabling more relevant, accurate, and context-aware AI responses

- A flexible LLM gateway to integrate preferred LLM and embedding providers, with optional hosting of embedding and ranking models in memory or through model-serving components for maximum operational flexibility



AgentOps and agent chassis capabilities: Telco IT AI Engine standardizes the design, orchestration, and operation of multi-agent systems at scale, with built-in capabilities for orchestration, memory management, observability, and guardrails:

- Multi-agent orchestration to build complex AI workflows by orchestrating specialized agents using standardized patterns (sequential, parallel, and supervisor flows) enabling sophisticated, multi-step automation
- Short and long term memory support for agents and conversational systems, enabling multi turn, context aware experiences for operators and customers
- Comprehensive observability stack powered by OpenTelemetry (OTel), providing unprecedented visibility into AI applications behaviour, performance, and quality metrics
- Safety guardrails to protect AI applications from harmful inputs and ensure safe, compliant responses aligned with telecom regulations and enterprise policies
- An anonymization library to help protect personal data across all GenAI, ensuring data privacy and regulatory compliance
- Agent chassis to accelerate delivery and remove risk, enabling teams to focus on logic rather than infrastructure plumbing, and standardizing how organization design, orchestrate, and operate multi agent systems
- Agent-to-agent (A2A) communication and model context protocol (MCP)-based access to tools and data through standardized, plug-and-play integrations, creating seamless ecosystems of agents



Platform architecture, security, and operations: Telco IT AI Engine is engineered as a robust, cloud agnostic platform with strong operational capabilities suitable for telco grade workloads:

- Cloud agnostic deployment with support for public cloud, private cloud, and air gapped on premise environments, and a deployment footprint that can be sized per use case and domain
- Backup and restore for key data stores to ensure resilience and operational continuity
- Backward compatibility framework for smooth upgrades of existing applications across Telco IT AI Engine releases, avoiding disruption to production pipelines
- Integrated security and access management that secures access and data protection across all systems, including encrypted data, trusted service connections, and Keycloak-based user authentication

Delivering business value: From AI pilots to production at scale

Ericsson Telco IT AI Engine delivers [business value](#) by operationalizing AI across OSS/BSS, IT, and network domains, addressing the technical and organizational barriers that have historically confined AI to siloed pilots:

- 1. Accelerate AI at scale across telco domains**
 - A standardized, enterprise-grade AI platform unifies MLOps, LLMOps, and AgentOps to rapidly design, deploy, and operate ML, GenAI, and agentic AI applications.
 - Reusable pipelines, prebuilt components, and low-code/autoML capabilities reduce duplication, avoid fragmented tooling, and accelerate high-impact use cases such as anomaly detection, forecasting, correlation, and intelligent assistants across markets and domains.
- 2. Improve AI trust, safety, and observability**
 - GenAI observability with OpenTelemetry based tracing and evaluations provide distributed tracing, quality scoring (relevance, accuracy, hallucination, toxicity), and real time dashboards, enabling continuous monitoring.
 - Robust guardrails detect prompt injection, jailbreak attempts, and harmful content, with extensible policies for telecom grade safety, regulatory compliance, and enterprise specific policies.
- 3. Protect privacy and ensure regulatory compliance**
 - An anonymization library automatically detects and anonymizes personally identifiable information (PII) in inputs and outputs, with configurable masking and replacement strategies tailored to regulatory requirements.
 - Data encryption at rest and in transit, mutual transport layer security (mTLS) protected service to service communication and integrated identity and access management (IAM) ensure robust security and privacy suitable for telco grade workloads.
- 4. Reduce total cost of ownership through platform standardization and cloud agnostic design**
 - Cloud agnostic deployment with support for on premises and air gapped installations ensures data sovereignty and alignment with regulatory and security requirements.
 - Platform services, including MLOps, LLMOps, AgentOps, observability, backup and restore, and security, minimize duplication and streamline operations across use-cases and domains.
- 5. Enable sophisticated agentic AI and autonomous operations**
 - Multi agent frameworks, memory, and agent chassis support design and operation of complex agentic systems for closed loop assurance, autonomous operations, and cross domain workflows.
- 6. Align with Ericsson data platforms and telco portfolio**
 - Out of the box integration with [Ericsson Telco DataOps Platform](#) and strong alignment to Ericsson OSS/BSS, IT, and network portfolios enables seamless, end-to-end solutions, from data to AI to action.

Operationalize AI at scale for intelligent telco operations

Ericsson Telco IT AI Engine enables autonomous operations, accelerated digitalization, and innovation at scale. Paired with Telco DataOps Platform for AI-ready data, the Telco IT AI Engine provides a scalable, cloud-native, and trusted platform with unified lifecycle management, open ecosystem integration, and low-code accelerators to rapidly build and operationalize ML, GenAI, and agentic AI use cases.

By addressing the fundamental barriers such as siloed data, fragmented platforms, and trust concerns, Ericsson Telco IT AI Engine transforms AI from a boutique pilot initiative into a strategic operational capability, enabling us to deliver autonomous, intelligent telco operations that drive competitive advantage, customer loyalty, and business growth.



Authors



Ramesh Samayamantula
Strategic Product
Manager - Data & AI



Indranil Roy Chowdhury
Chief AI-Data Architect



Pooja Singh
Strategic Product Manager
– Data & AI



Mohit Bhargava
OSS/BSS Product
Marketing Manager

Further reading

[AI-powered telecom data and analytics solutions](#)

[Ericsson telco IT AI Apps](#)

[Ericsson Telco DataOps Platform](#)

[How to accelerate automation in OSS/BSS with agent fabric](#)

[From data to decisions: Make agentic AI-driven telecom operations a reality](#)

[How multi-agent AI is transforming telco product configuration](#)

[Ericsson accelerates AI innovation and industrialization for OSS/BSS](#)

[Leverage agentic AI for intelligent charging and billing operations](#)

[When will AI agents begin to design and deploy new AI apps for OSS/BSS?](#)

[Maximize OSS/BSS impact with AI and Gen-AI](#)

[Business Value Pathways](#)

About Ericsson

Ericsson's high-performing networks provide connectivity for billions of people every day. For nearly 150 years, we've been pioneers in creating technology for communication. We offer mobile communication and connectivity solutions for service providers and enterprises. Together with our customers and partners, we make the digital world of tomorrow a reality.

www.ericsson.com