CLOUD-NATIVE NEXT CHAPTER –

AGENTIC AI-BASED OPERATING MODELS

V1.0

ngmn.org

WE MAKE BETTER CONNECTIONS

CLOUD-NATIVE NEXT CHAPTER – AGENTIC AI-BASED OPERATING MODELS

by NGMN Alliance

Version:	1.0
Date:	24 March 2026
Document Type:	Public
Programme:	Mastering the Route to Disaggregation
Approved by / Date:	NGMN Board, 16 March 2026

The information contained in this document represents the current view held by NGMN Alliance e.V. on the issues discussed as of the date of publication. This document is provided "as is" with no warranties whatsoever including any warranty of merchantability, non-infringement, or fitness for any particular purpose. All liability (including liability for infringement of any property rights) relating to the use of information in this document is disclaimed. No license, express or implied, to any intellectual property rights are granted herein. This document is distributed for informational purposes only and is subject to change without notice. Readers should not design products based on this document.

CONTENTS

EXECUTIVE SUMMARY

This document defines the NGMN view on the roadmap for evolving cloud-native telecom operations by integrating advanced artificial intelligence (AI) technologies. Building on the NGMN Cloud-Native Manifesto and on the Cloud-Native Computing Foundation (CNCF) assessment frameworks (CNMM), it provides practical guidance for mobile network operators, to adopt GenAI-powered operating models, drive tangible business outcomes, and elevate operational maturity in people, processes, and technology. Five "AI adoption levels" are defined, leveraging CNMM "maturity levels" as a baseline for the introduction of AI in the network operation. This document covers readiness assessment, phased adoption, tooling and frameworks, best practices, operational use cases, and organisational transformation. The content emphasises industry collaboration, responsible AI governance, skill development, and continuous optimisation to support scalable, resilient, and autonomous networks.

For clarity, the terms AI, GenAI, and Agentic AI used throughout refer to the same evolving family of technologies — each representing increasing sophistication and autonomy. All principles, strategies, and recommendations in this document apply broadly to advanced AI in general, including GenAI and Agentic AI systems.

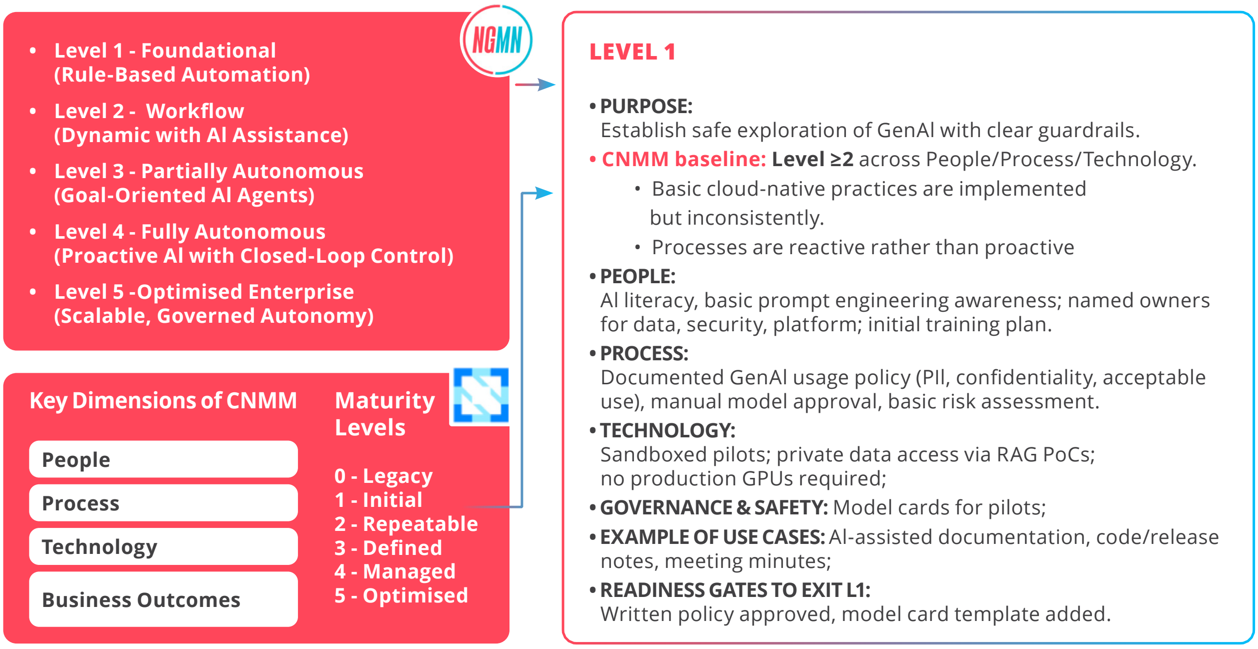
01 INTRODUCTION

With this document, NGMN provides guidance to mobile operators for the adoption of AI to operate the network, providing an adoption path with respect to their cloud-native maturity. To achieve AI proficiency, the proposed framework defines five levels of AI adoption, each mapped to the corresponding CNMM maturity level defined by CNCF.

This publication evaluates the alignment of the "NGMN Cloud-Native Manifesto" with CNCF cloud-native frameworks, emphasising telecom-specific adaptations across technology, processes, and skills essential for transformation. It leverages the established Cloud-Native Maturity Model (CNMM), published by CNCF, to define clear maturity levels, guiding operators from initial cloud-native adoption to fully optimised states.

Importantly, this document maps those cloud-native maturity levels to corresponding stages of AI readiness, outlining how artificial intelligence—including Generative AI (GenAI) and its more autonomous form, Agentic AI—can be progressively integrated into telecom operating models. This phased approach supports a structured transition from early AI experiments through standardised AI-driven workflows toward fully Agentic AI-enabled autonomous network operations. Cloud-native model adoption also depends heavily on the ecosystem and available commercial models offered by vendors, supporting and driving openness and agility in using their products and services. Generally, vendors still offer a vertically integrated cost model that does not align with cloud-native best practices envisioning horizontal cloud and tools.

The guidance detailed within covers practical implementation of AI-powered DevOps, predictive AIOps, customer-centric observability, and ethical AI practices, while maintaining regulatory compliance. The publication explores key GenAI use cases—like AI-assisted documentation, architecture optimisation, autonomous operational management—and demonstrates how telecom operations can evolve into self-healing networks with dynamic resource allocation and predictive maintenance powered by AI. Emphasising the importance of cross-industry collaboration, this work establishes best practices for building scalable, self-optimising networks powered by Agentic AI, leveraging cloud-native principles to enhance operational agility and innovation. Through this comprehensive approach, mobile network operators can accelerate their transformation journey towards intelligent, autonomous, and resilient telecom ecosystems.



- **Level 1 - Foundational (Rule-Based Automation)**
- **Level 2 - Workflow (Dynamic with AI Assistance)**
- **Level 3 - Partially Autonomous (Goal-Oriented AI Agents)**
- **Level 4 - Fully Autonomous (Proactive AI with Closed-Loop Control)**
- **Level 5 -Optimised Enterprise (Scalable, Governed Autonomy)**

Key Dimensions of CNMM	Maturity Levels
People	0 - Legacy 1 - Initial 2 - Repeatable 3 - Defined 4 - Managed 5 - Optimised
Process	
Technology	
Business Outcomes	

LEVEL 1

- **PURPOSE:** Establish safe exploration of GenAI with clear guardrails.
- **CNMM baseline: Level ≥2** across People/Process/Technology.
  - Basic cloud-native practices are implemented but inconsistently.
  - Processes are reactive rather than proactive
- **PEOPLE:** AI literacy, basic prompt engineering awareness; named owners for data, security, platform; initial training plan.
- **PROCESS:** Documented GenAI usage policy (PII, confidentiality, acceptable use), manual model approval, basic risk assessment.
- **TECHNOLOGY:** Sandboxed pilots; private data access via RAG PoCs; no production GPUs required;
- **GOVERNANCE & SAFETY:** Model cards for pilots;
- **EXAMPLE OF USE CASES:** AI-assisted documentation, code/release notes, meeting minutes;
- **READINESS GATES TO EXIT L1:** Written policy approved, model card template added.

Figure 1: Visualisation of the five "AI adoption levels" defined by NGMN leveraging CNCF's CNMM "maturity levels" as a baseline for the introduction of AI in network operation

02 CLOUD NATIVE MATURITY MODEL ASSESSMENT

Cloud-Native Maturity Model (CNMM) by CNCF is a structured framework designed for telecom operators to evaluate their maturity in adopting cloud-native principles.

Based on the Capability Maturity Model (CMM), it defines five levels of maturity and integrates business outcomes with three critical dimensions: People, Process, and Technology. In this chapter, NGMN assesses this methodology to relate it to the adoption of Agentic AI for the operation of the network.

2.1 OBJECTIVES OF CNMM

1. Identify current cloud-native maturity level.
2. Define target maturity levels aligned with strategic goals.
3. Develop a transformation roadmap for cloud-native excellence.

Building on this foundational cloud-native maturity, NGMN proposes, with this document, the integration of Agentic AI–based frameworks as the next evolutionary step for telecom operating models. By leveraging the maturity levels and dimensions identified in CNMM, operators can assess their readiness for embedding Agentic AI capabilities—enabling intelligent autonomy, proactive orchestration, and closed-loop operational workflows.

This mapping ensures that AI-driven transformations align with organisational goals, while maintaining governance, security, and compliance, thus bridging cloud-native excellence with advanced, autonomous network operations.

2.2 KEY DIMENSIONS OF CNMM

1. PEOPLE:
Focuses on skills, culture, leadership, security, and training.

- **Sub-dimensions:**
  Skills & Expertise, Culture & Collaboration, Leadership, Security Team, Training.

2. PROCESS:
Evaluates workflows, governance, automation, and incident management.

- **Sub-dimensions:**
  Workflow Standardisation, Automation, Governance & Compliance, Monitoring & Incident Management.

3. TECHNOLOGY:
Assesses cloud-native platforms, architecture, and integration.

- **Sub-dimensions:**
  Cloud-native Platform, Architecture, Cloud Integration, Security Tools.

4. BUSINESS OUTCOMES:
Measures alignment with business goals such as cost optimisation, network efficiency, customer experience, time-to-market, and regulatory compliance.

2.3 MATURITY LEVELS

CNMM defines five maturity levels based on CMM (Capability Maturity Model):

- **0 - Legacy:**
  No cloud-native adoption.
- **1 - Initial (Ad-hoc):**
  Minimal understanding; unstructured adoption.
- **2 - Repeatable (Opportunistic):**
  Basic practices implemented inconsistently.

- **3 - Defined (Systematic):** Standardised practices aligned with business goals.
- **4 - Managed (Measured):** Proactive monitoring and optimisation.
- **5 - Optimised:** Continuous improvement and innovation.

2.4 ASSESSMENT AREAS

PEOPLE ASSESSMENT:

- **Skills & Expertise:** Evaluate technical knowledge (e.g., Kubernetes, CI/CD).
- **Culture & Collaboration:** Adoption of DevOps and cross-functional teamwork.
- **Leadership:** Commitment to cloud-native initiatives.
- **Security Team:** Expertise in cloud-native security practices.
- **Training:** Availability of training programmes and certifications.

PROCESS ASSESSMENT:

- **Workflow Standardisation:** Consistency and documentation of processes.
- **Automation:** Use of CI/CD pipelines, GitOps, and Infrastructure-as-Code (IaC).
- **Governance & Compliance:** Integration of compliance checks into workflows.
- **Monitoring & Incident Management:** Implementation of observability tools (e.g., Prometheus, Grafana).

TECHNOLOGY ASSESSMENT:

- **Cloud-native Platform:** Adoption of containerisation, orchestration, and automation tools.
- **Architecture:** Use of microservices, Service Mesh, and APIs.
- **Cloud Integration:** Integration of legacy systems with cloud-native platforms.
- **Security Tools:** Implementation of tools for vulnerability scanning, runtime protection, and compliance.

BUSINESS OUTCOMES:

- **Cost Optimisation:** Reduction in operational costs through automation.
- **Network Efficiency:** Improved scalability and resource utilisation.
- **Customer Experience:** Enhanced service quality and reliability.
- **Time-to-Market:** Faster deployment of new services.
- **Regulatory Compliance:** Proactive risk management and automated compliance.

2.5 RESULTS OF CNMM AND AGENTIC AI ADOPTION

The self-assessment provides:

1. **Current Maturity Level:** Understanding of the organisation's current state.
2. **Target Maturity Level:** Goals based on business priorities.
3. **Gap Analysis:** Identification of improvement areas.
4. **Actionable Insights:** Recommendations for achieving target maturity.

CNMM as foundations of Agentic AI-based Operating Model:

CNMM is a comprehensive tool for telecom operators to evaluate and enhance their cloud-native maturity. It aligns technical capabilities with business outcomes, ensuring a structured transformation journey.

- Cloud-native adoption requires cultural, operational, and technical transformation.
- Business outcomes must drive cloud-native initiatives.

- The methodology provides a roadmap for achieving agility, scalability, and innovation.

Building on these maturity levels and transformation principles, Agentic AI–based operating models represent the next evolution in telecom operations. By mapping cloud-native maturity stages to AI readiness, operators can systematically plan their transition toward autonomous, intelligent network management.

This approach embeds Agentic AI capabilities into existing people, process, and technology frameworks to enhance decision-making, automation, and resilience — while maintaining governance, security, and compliance.

For a mobile operator, being aware of its maturity level is not only a foundation for cloud-native excellence but it is also a strategic guide for integrating cutting-edge Agentic AI to realise fully autonomous, next-generation telecom networks.

03 AGENTIC AI-BASED OPERATING MODELS

3.1 LEVEL OF CLOUD-NATIVE MATURITY REQUIRED FOR AGENTIC AI-BASED OPERATING MODELS

To effectively integrate Agentic AI into telecom operating models, it is essential to first evaluate the level of cloud-native maturity. This assessment determines readiness and provides motivation for adopting Agentic AI within cloud-native networks. Understanding how maturity correlates with AI integration helps identify relevant use cases, drivers, opportunities, and implications critical for a successful transition.

3.1.1 Motivation for Integrating GenAI in Cloud-Native Networks

DRIVERS:

- **Operational complexity:**
  The presence of cloud-native functions (CNFs), multi-vendor environments, and heterogeneity across edge and core networks make manual operations unsustainable.
- **Time to restore and skills gap:**
  AI copilots help reduce Mean Time To Repair (MTTR) and augment scarce expertise.
- **Data deluge:**
  Extensive telemetry, logs, traces, and configuration data provide rich inputs for retrieval and reasoning.
- **Cost and sustainability:**
  Enables right-sizing, token efficiency, and carbon-aware scheduling.

OPPORTUNITIES:

- **AIOps and closed-loop automation:**
  Facilitates predictive detection, root cause analysis, and safe automated remediation.
- **Customer care transformation:**
  Retrieval-Augmented Generation (RAG) over telecom knowledge reduces Average Handle Time (AHT) and improves Customer Satisfaction (CSAT).
- **Network engineering:**
  AI-assisted design, capacity planning, and digital twin validation support smarter network evolution.
- **New business models:**
  Includes AI-as-a-Service (AIaaS) for partners, AI-optimised network slicing, and developer platforms powered by Agentic AI.

IMPLICATIONS:

- **Data governance:**
  Ensures lineage, data sovereignty, retention policies, and access controls for AI prompts and embeddings.
- **Security & safety:**
  Addresses model supply chain integrity, defences against prompt injection, and output guardrails.
- **Organisation & skills:**
  Creates new roles such as LLMOps, AI Ops Engineer, Prompt Engineer, and Network Data Scientist, while incorporating Human-In-The-Loop (HITL) workflows.
- **Vendor expectations:**
  Emphasises openness in APIs and telemetry, safety Service Level Agreements (SLAs), model cards, and attestations.

3.1.2 Quick Readiness Checks per Level

Level 1 – Foundational (Rule-Based Automation)

- AI adoption begins with basic, rule-driven robotic process automation (RPA) where actions and sequences are predefined. Governance is established with policies, PII controls, and sandboxed pilots. Example: Automating invoice data extraction.

Level 2 – Workflow (Dynamic with AI Assistance)

Actions remain defined, but sequences are dynamically managed via routers or LLMs, enabling branching workflows. Model registries, basic evaluations, GitOps, and early SLOs guide initial production AI use. Example: Customer email drafting or Retrieval-Augmented Generation.

Level 3 – Partially Autonomous (Goal-Oriented AI Agents)

AI agents plan and adapt task sequences with minimal human oversight using domain-specific toolkits. Standardised pipelines, lineage tracking, drift detection, red-team exercises, and enforced SLOs support operational reliability. Example: Autonomous multi-system ticket resolution.

Level 4 – Fully Autonomous (Proactive AI with Closed-Loop Control)

AI systems operate independently across domains, setting goals and adapting to outcomes, with closed-loop automation and rollback safeguards. Advanced GPU scheduling, confidential computing, and carbon-aware telemetry enhance efficiency and compliance. Example: AI-driven strategic research and synthesis.

Level 5 – Optimised Enterprise (Scalable, Governed Autonomy)

Enterprise-wide governance with policy-as code, audits, attestations, and global resilience ensures continuous optimisation. KPIs monitor performance, cost, security, and sustainability at scale. Example: Fully autonomous Agentic AI platforms driving next-gen telecom operations.

3.1.3 Mapping to AI Adoption

- **People:**
  Add AI roles/skills, HITL oversight, red-team readiness, Responsible AI literacy at each level.

- **Process:**
  Extend Workflow Standardisation and Automation with LLMOps gates; embed Governance & Compliance with model cards, lineage, fairness/bias evals; enhance Incident Management to include model/inference SLOs and drift.

- **Technology:**
  Extend Platform with AI serving, vector/RAG, GPU scheduling, observability for LLMs; Security Tools with model artifact signing, policy-as-code for ML, confidential computing.

- **Business Outcomes:**
  Add model quality/safety, inference SLOs, cost/tokens and CO2e, CSAT/AHT, network efficiency uplift from AI.

Below is a compact, telco-focused five-level maturity model for GenAI-based operating models, aligned to CNMM (People, Process, Technology, Business Outcomes) and informed by the "Cloud-Native Next Chapter — Agentic AI-based Operating Models". It defines what "good" looks like at each level, the minimum cloud-native baseline required, gating controls, and examples of telco use cases unlocked as per defined levels. It assumes Level 0 (Legacy) is pre-GenAI and starts at Level 1.

3.2 AGENTIC AI-BASED OPERATING MODELS MATURITY MAPPING TO CNMM

Level 1 - Foundational (Aware, Safe pilots)

- **Purpose:**
  Establish safe exploration of GenAI with clear guardrails.

- **CNMM baseline:**
  Level ≥2 across People/Process/Technology.

  > Basic cloud-native practices are implemented but inconsistently.

  > Initial adoption of CI/CD Pipeline Tools, containerisation, Kubernetes, GitOps Tool (FluxCD or ArgoCD), basic monitoring & observability tools.

  > Processes are reactive rather than proactive based on automation with DevOps and GitOps operating model.

- **People:**
  AI literacy, basic prompt engineering awareness; named owners for data, security, platform; initial training plan.

- **Process:**
  Documented GenAI usage policy (PII, confidentiality, acceptable use), manual model approval, basic risk assessment.
- **Technology:**
  Sandboxed pilots; private data access via RAG PoCs; no production GPUs required; secured access to LLM APIs or a minimal self-hosted small model; basic logging.
- **Governance & Safety:**
  Model cards for pilots; PII redaction; output filtering; HITL.
- **Example of use cases:**
  AI-assisted documentation, code/release notes, meeting minutes, runbook drafting, ticket summarisation.
- **Outcomes:**
  Faster knowledge access; measurable PoC throughput; zero data leakage incidents.
- **Readiness gates to exit L1:**
  Written policy approved, model card template adopted, PII scanning in place, pilot review checklist defined.

Level 2: Enabled (Controlled pilots, initial production)

- **Purpose:** Move from ad hoc PoCs to controlled early production with basic LLMOps.
- **CNMM baseline: Level ≥3** across People/Process/Technology.
  > Cloud-native practices are standardised across teams, DevSecOps and GitOps principles are widely adopted.

  > Clear alignment between technical initiatives and business outcomes, Kubernetes is used with Helm Charts or Yaml Manifests, CI/CD pipelines are implemented, GitOps workflows are widely adopted, Security tools are integrated into CI/CD pipeline, and monitoring and observability tools (Prometheus/Grafana, OpenTelemetry) are actively used for centralised metrics, logs, and traces.

  > Processes are documented and repeatable, automation is implemented for most workflows, and use of advanced CI/CD pipelines with GitOps tools.
- **People:**

  Cross-functional pod (platform/SRE, ML/data, security); defined RACI for model life-cycle; basic red-team playbook.
- **Process:**

  Initial MLOps/LLMOps pipeline (data validation → evaluation → approval → deployment); incident playbook for AI issues; GitOps for configurations/prompts.
- **Technology:**

  Model registry; basic KServe (or equivalent) for serving; vector DB for internal RAG; initial GPU nodes (or CPU-serving for small/optimised models); observability for p50/p95 latency and cost/tokens.
- **Governance & Safety:**

  Manual approvals; offline eval suites; content/PII guardrails; audit trail of model/prompt versions.
- **Example of use cases:**

  NOC copilot (ticket triage/summarisation), RAG over network SOPs and vendor docs, change request drafting, business reporting Q&A.
- **Outcomes:**

  10–20% MTTR reduction on targeted incidents; <X ms p95 latency for internal copilots; cost per 1K tokens tracked.
- **Readiness gates to exit L2:**

  Registry + signed artefacts; evaluation reports required for go-live; SLOs defined (latency, quality); basic runtime observability.

Level 3 — Integrated (Production LLMOps, SLO-backed services)

- **Purpose:** Standardise LLMOps, integrate GenAI into operational workflows with SLOs.
- **CNMM baseline: Level 3–4** across People/Process/Technology.
  > Cloud-native practices are standardised across teams or highly skilled advanced cloud-native practices, strong collaboration between technical and business teams, SRE practices are in place to ensure reliability and performance, and security is fully integrated into cloud-native workflows and managed.

> Clear alignment between technical initiatives and business outcomes, advanced and automated tool usage for scalability, resilience, and operational efficiency (Kubernetes clusters optimised, GitOps Tools for Continuous deployment and rollback, advanced observability tools for real-time insights and anomaly detection).

> Processes are documented and repeatable across teams, are optimised for efficiency and scalability, and Incident response is proactive with root cause analysis and prevention mechanisms.

- **People:**

  SRE practices expanded to inference workloads (SLIs/SLOs for latency/throughput); platform team enabling self-service patterns; product ownership for GenAI services.

- **Process:**

  Standard pipelines with canary/shadow testing; dataset/version lineage; scheduled re-evaluation; AI incident postmortems; change management via GitOps (infra/app/model/prompt).

- **Technology:**

  KServe (or similar) for rollout strategies; Kubeflow/Ray for training/evals; Kueue/Volcano for batch/gang scheduling; service mesh with mTLS; vector DB in prod; OpenTelemetry + early OpenLLMetry; autoscaling (KEDA Kubernetes Event-Driven Autoscaling / HPA Horizontal Pod Autoscaler).

- **Governance & Safety:**

  Policy as code gates in CI/CD/ML (quality, bias, safety); red team exercises; data access controls for prompts/embeddings; model cards mandatory.

- **Example of use cases:**

  Customer care GenAI with governed RAG; network planning assistant; change risk assessment with human approval; closed-loop suggestions (HITL).

- **Outcomes:**

  p95 latency SLOs met in prod; 20–30% ticket handling time reduction; accuracy/quality KPIs tracked; reproducible lineage for audits.

- **Readiness gates to exit L3:**

  SLOs reliably met across two+ services; automated rollbacks; drift detection live; model governance board operational.

Level 4 — Closed-Loop (Autonomous-1, safety-gated automation)

- **Purpose:** Introduce safe autonomy in well-bounded loops; scale multi-tenant, efficient inference/training.
- **CNMM baseline: Level ≥4** across People/Process/Technology.

  > Teams are highly skilled in advanced cloud-native practices, evidence of collaboration between DevSecOps, business, and platform teams, and Setting up SRE roles and practices.

  > Advanced and automated tool usage for scalability, resilience, and operational efficiency (Kubernetes self-healing configuration, advanced GitOps, advanced observability for real-time insight and anomaly detection, advanced security tools for proactive threat detection and runtime protection).

  > Processes are optimised for efficiency and scalability, continuous improvement mechanisms exist, evidence of process optimisation through feedback loops, and Incident response times are reduced, and root cause analysis is documented systematically.

- **People:**

  Ops/engineering adopt "automation-first"; clear exception handling; FinOps + sustainability roles engaged.

- **Process:**

  Continuous eval with synthetic traffic; drift/guardrail breaches trigger quarantine/rollback; CAB integrates model risk scoring; periodic fairness/robustness reviews.

- **Technology:**

  Advanced accelerator management (MIG/

MPS, quotas); Dynamic Resource Allocation, bin-packing, pre-emption; multi-cluster placement; model gateway and request shaping; confidential computing pilots for sensitive data; carbon/cost telemetry; policy-driven routing (canary/AB/shadow) at scale. Digital Twin for safe and pre-empted automation.

- **Governance & Safety:**

  Auto-gated promotions; explainability where required; DPIA (Data Protection Impact Assessment) where needed; vendor accountability SLAs for safety fixes.

- **Example of use cases:**

  Self-healing auto-remediation for low-risk incidents; proactive capacity scaling; digital twin-driven change validation; RIC xApps/rApps with LLM-assisted policy recommendations (HITL).

- **Outcomes:**

  30–50% MTTR reduction; 15–25% GPU utilisation uplift; controlled auto-remediation with zero critical incidents; CO2e per job tracked and trending down.

- **Readiness gates to exit L4:**

  Documented set of closed loops with risk taxonomy; zero P1 safety incidents over N quarters; confidential pipelines validated; carbon/cost-aware scheduling in pilot.

Level 5 — Optimised (Responsible autonomy at scale)

- **Purpose:**

  Enterprise-wide, audited, sustainable GenAI with federated governance and continuous optimisation.

- **CNMM baseline: Level 5** across People/Process/Technology.

  > Continuous improvement and innovation driven by cloud-native principles, Teams are empowered to innovate and experiment with new technologies, continuously refine SRE practices based on feedback and metrics from monitoring & observability, and AIOps systems. Focus on proactive optimisation and efficiency improvements.

  > Cloud-native and AI to support Autonomous network, fully automated, predictive, and AI/ML-driven tool usage, fully automated advanced CI/CD pipelines with manual intervention, and zero-downtime deployments are achieved, predictive monitoring and observability tools provide actionable insights, security tools are fully automated and integrated into all workflows, and security incidents are automatically detected and mitigated in real-time.

  > Processes are continuously improved based on business and operational feedback, are documented and tracked, predictive monitoring and AI-driven incident prevention are implemented, and predictive analytics tools (e.g., AI/ML) are used for proactive incident prevention.

  > Business outcomes are consistently achieved and measured.

- **People:**

  Federated enablement model; continuous upskilling; contributions to open standards; well-practised human oversight.

- **Process:**

  Enterprise policy-as-code; continuous compliance (GDPR/data sovereignty, telecommunications regulations); external audits; post-incident learning institutionalised.

- **Technology:**

  Multi-region/global failover for model services; SLA-aware, cost/carbon-aware scheduling; advanced model optimisation (distillation, Mixture of Experts, caching); standardised confidential AI blueprints; unified metadata/catalogue. Digital Twin for simulation and validation of complex changes supporting new business use cases.

- **Governance & Safety:**

  Trust & Safety KPIs embedded; third-party attestations (provenance, SBOM/OCI artefacts); red-team and chaos-LLM exercises.

- **Example of use case:**

  AI-optimised network slices; autonomous operations with human oversight; GenAI at scale for customers and internal users; AIaaS for partners.

- **Outcomes:**

  Predictable ROI; audited compliance; continuous cost/CO2e reductions; high developer/analyst productivity.
- **Exit is ongoing optimisation:** maturity focuses on resilience, trust, sustainability, and ecosystem value.

3.3 TOOLS AND FRAMEWORK REQUIRED TO ENABLE GENAI-BASED OPERATING MODELS

AI, GenAI, and Agentic AI are driving fundamental changes in the design and management of cloud-native network environments. An AI-based operating model establishes a structured framework for embedding artificial intelligence into the management and operational processes of telecom networks. This model clearly delineates roles, responsibilities, workflows, and the technological enablers needed for leveraging advanced AI for decision-making, automation, and continuous optimisation.

Core elements of this framework include data-driven intelligence, AI-powered orchestration, predictive and prescriptive analytics, and closed-loop operational feedback. These capabilities enable networks to achieve higher adaptability, resilience, and efficiency while maintaining governance, regulatory compliance, and enterprise-grade scalability.

Integrating AI in a cloud-native context unlocks a range of use cases and operational scenarios—driven by the need to address complexity, accelerate troubleshooting, automate service management, and deliver actionable insights at scale.

The integration of Low-Code and No-Code development approaches can significantly accelerate AI adoption in telecom by enabling intuitive, natural language-driven workflow creation and democratising access to advanced automation capabilities.

These tools and technologies ultimately empower telecom providers to innovate, manage risk, optimise costs, and respond swiftly to changing demands by harnessing the full capabilities of modern AI across every layer of digital network operations.

3.3.1 Foundational Components

- **Data Architecture:**

  Robust and scalable data infrastructure is essential to manage both structured and unstructured data, enabling AI readiness across the network.
- **Digital Twin:**

  Virtual representations of network environments support simulation, testing, and validation of AI-driven operational decisions.
- **AI Flow Building Tools:**

  Platforms and frameworks that facilitate the creation, management, and orchestration of AI workflows and tasks.
- **Organisational Skillsets:**

  Cross-functional teams with expertise in AI/ML, data engineering, network operations, and AI governance are critical for successful AI adoption.

3.3.2 Infrastructure and Platform

- **Cloud-Native Infrastructure:**

  Kubernetes, OpenShift, and K3s orchestrate containers and scale AI services efficiently.
- **Network Functions Life-Cycle Management:**

  CI/CD pipelines and Nephio automate life-cycle management of virtual and cloud-native network functions.
- **Service Mesh:**

  Istio and Linkerd provide secure, observable, and controlled communication between service components.
- **MLOps/AIOps Platforms:**

  Kubeflow, MLflow, and Argo automate AI model training, deployment, and monitoring.

3.3.3 AI Model and Development Platforms

- **LLM Platforms:**

  OpenAI, Hugging Face, and Meta Llama provide generative AI capabilities for automation and informed decision-making.

- **Multimodal Frameworks:**

  LangChain, LlamaIndex, and Haystack enable knowledge-driven, retrieval-augmented AI applications.

- **AI Agent Frameworks:**

  AutoGen, CrewAI, OpenDevin support autonomous task decomposition and operational workflows.

- **Reinforcement Learning:**

  Ray RLlib, Stable-Baselines3 facilitate dynamic resource allocation and policy learning optimisation.

3.3.4 Data and Observability Tools

- Data schema standardisation across domains and standards

- **Data Pipelines:**

  Apache Kafka, Spark, Airflow, and Feast manage data ingestion, processing, and feature engineering.

- **Monitoring and Telemetry:**

  Prometheus, Grafana, and OpenTelemetry collect and visualise network performance metrics.

- **Vector and Knowledge Databases:**

  FAISS, Weaviate, and Milvus enable semantic search and advanced AI reasoning capabilities.

04 EVOLVING ROLES IN ORGANISATIONS WITH AI-BASED OPERATING MODELS

AI will change everyone's role in the organisation, and this section outlines a few examples to explore this transition.

4.1 IMPACT OF AI ON ORGANISATIONAL ROLES

- **Network Operations Automation:**

  Automated fault detection, traffic anomaly detection, predictive network management, root-cause analysis and post-incident learning. The role of network engineers and operations engineers is shifting and will evolve. Traditional NOC roles will be partially replaced by AI platforms.

- **Security Protection:**

  Real-time threat detection and anomalous behaviour pattern recognition. Security engineers will develop security models created from AI and applied for AI traffic. Their role will evolve.

- **Automated Testing and Capacity Forecasting:**

  Digital twin–based network simulation for safe testing and performance forecasting, capacity engineers' and DevOps engineers' roles will shift and evolve, leveraging proactive intelligent resource management and dynamic allocation to optimise performance and ensure sustainable operations across multi-cloud and edge environments.

- **Network Architecture and Design:**

  Increased Complexity of 5G/6G and cloud-native networks: Manual management becomes unsustainable. Architects will use AI to develop new services, relying on AI-powered "what-if analysis" and digital twin simulations to evaluate network design alternatives, assess risks, and validate changes before implementation.

- Cloud Templates Creating and Hydration:

  We can use GenAI-based code development to build templates like cloud deployments. DevOps engineers will shift.

- **Project Management and Tracking:**

  Meeting minutes and project milestones are already being created with AI. PM would still be valuable in connecting emotional intelligence, navigating complex stakeholder alignment and driving vision forward.

- **New Services and Business Cases:**

  AI can develop strong business cases by numbers analysis and ROI calculations. Supporting and evolving the role of business teams.

- **Strategy and Leadership:**

  Analysing market trends, global adoption of technologies and ROI publications, AI can give inputs to executives in strategy and vision planning. CTO role would also evolve in the new world.

4.2 SHIFTING PEOPLE, PROCESS AND TECHNOLOGY IN AI-BASED OPERATING MODELS

The adoption of AI—ranging from foundational AI through GenAI to advanced Agentic AI—profoundly transforms telecom organisations at every level. This section explores the critical shifts required in people's roles and mindsets, the redesign of operational workflows, and the broader organisational changes needed to fully realise AI's potential.

4.2.1 People: Shifting Roles and Mindsets

- From manual operators to AI-augmented decision-makers.
- Emergence of new roles (e.g., AI Ops Engineers, Prompt Engineers, AI Trainers).
- Redefining accountability in AI-influenced decisions.
- Fostering a culture of continuous learning and AI trust.

4.2.2 Process: Redesigning Operational Workflows

- Moving from static playbooks to adaptive, AI-driven automation.
- Integrating GenAI into CI/CD, incident response, and capacity planning.
- Governance and human-in-the-loop (HITL) process design.
- Managing AI model life-cycle within operational processes.

4.2.3 Organisation: Organisational Impact and Change Management

- Restructuring teams around AI-driven operations.
- Addressing resistance to AI-led change.
- Change management frameworks for adoption.
- What will make this model successful? How can mobile network operators get ready for this?

05 CONCLUSIONS

The adoption of Agentic AI represents the next chapter in network operations, building on the investments made and the maturity achieved in cloud-native adoption. This evolution requires a clear roadmap to further transform cloud-native telecom operations through the integration of advanced artificial intelligence (AI) technologies. It is not only a technological shift, but also a cultural transformation that spans the entire organisation. To successfully adopt AI, the following key success factors and actions must be considered.

5.1 KEY SUCCESS FACTORS

1. **Well-Defined Operating Framework and Human-AI Role Clarity**

   A successful GenAI model clearly delineates the roles and responsibilities between AI and human operators. AI is responsible for augmenting decision-making, automating repetitive tasks, and offering predictive insights, while humans oversee strategy, governance, exception handling, and AI supervision.

2. **Robust Data and Infrastructure Readiness**

   GenAI performance hinges on high-quality, real-time data from across the network — including event logs, traffic patterns, behavioural telemetry, and security signals. A cloud-native infrastructure that supports seamless data flow, storage, and integration is essential for success.

3. **Strong AI Governance and Trust Mechanisms**

   As AI systems assume greater responsibility, robust governance structures must be in place — ensuring model explainability, compliance with operational policies, ethical standards, and clear audit trails for AI-driven decisions.

To ensure the success of a GenAI-based operating model in cloud-native network environments, mobile network operators must move beyond viewing GenAI as a stand-alone tool and instead embrace it as a core enabler of operational transformation. This model defines a structured framework where generative AI is embedded into network management and operations, enabling smarter decision-making, adaptive automation, and continuous optimisation through data-driven intelligence, AI-assisted orchestration, and closed-loop operational feedback.

5.2 CALL TO ACTION FOR CSPS AND INDUSTRY

Develop AI-Ready Teams for Network Operations

- Build cross-functional teams including AI/ML engineers, LLMOps specialists, network operations engineers, and HITL (Human-in-the-Loop) monitors.
- Foster Responsible AI literacy: understand bias, ethics, compliance, and security risks in AI-driven network automation.

Integrate GenAI into Cloud-Native Network Workflows

- Integrate LLMOps gates into cloud-native CI/CD and network automation pipelines to ensure all AI-driven decisions are verified.

Upgrade Cloud-Native Platforms for GenAI

- Deploy AI serving, vector databases/RAG, GPU scheduling, and LLM observability to support large-scale intelligent network operations.

Measure AI-Driven Network Outcomes

- Continuously monitor AI impact on network resource allocation, traffic prediction, and fault response to validate ROI.

Collaborate Across Industry and Standards

- Collaborate closely with industrial alliances (such as NGMN), cloud providers, network vendors, and SDOs to create a unified, scalable, and interoperable GenAI-driven cloud-native network ecosystem.
- Share best practices to promote standardisation of AI-driven network operations across CSPs, reducing deployment and operational risks.

06 LIST OF ABBREVIATIONS

AHT	Average Handle Time
API	Application Programming Interface
CI/CD	Continuous Integration / Continuous Deployment
CAB	Change Advisory Board
CMM	Capability Maturity Model
CNMM	Cloud-native Maturity Model
CNCF	Cloud-native Computing Function
CNF	Cloud-native Network Function
CSAT	Customer Satisfaction (Score)
DRA	Dynamic Resource Allocation
HITL	Human-In-The-Loop
HPA	Horizontal Pod Autoscaler
KEDA	Kubernetes Event-Driven Autoscaling
KPI	Key Performance Indicator
LLM	Large Language Model
MIG	Multi-Instance GPU

MPS	Multi-Process Service
ML	Machine Learning
MTTR	Mean Time to Repair
NOC	Network Operation Centre
PII	Personally Identifiable Information
PNF	Physical Network Function
OCI	Open Container Initiative
RACI	Responsible Accountable Consulted Informed
RAG	Retrieval-Augmented Generation
RIC	RAN Intelligent Controller
SBOM	Software Bill of Materials
SLI	Service Level Indicator
SLO	Service Level Objective
SLA	Service Level Agreement
SRE	Site Reliability Engineer
VNF	Virtualised Network Function

07 REFERENCES

[1] https://www.cncf.io

[2] https://maturitymodel.cncf.io

08 FIGURES

ACKNOWLEDGEMENTS

TELUS, Sana Tariq	Project Lead
TIM, Fabrizio Moggio	Project Co-Lead
NGMN, Sparsh Singhal	NGMN Programme Manager

Special thanks to the following organisations and people for their contributions:

BT, Milan Lalovic	**Orange,** Dominique Vo
Chunghwa Telecom, Yuan-Mao Hung	**Telefónica,** Jesús Martín
Orange, Cédric Ollivier	

NEXT GENERATION MOBILE NETWORKS ALLIANCE

NGMN - Next Generation Mobile Networks Alliance - is a global, operator-driven organisation established by leading international mobile network operators (MNOs). As a global alliance of operators, vendors, and academia, NGMN provides industry guidance to enable innovative, sustainable and affordable next-generation mobile network infrastructure.

Key focus areas include Mastering the Route to Disaggregation, Green Future Networks, and 6G, while supporting the full implementation of 5G. NGMN drives global alignment of technology standards, fosters collaboration with industry organisations and ensures efficient, project-driven processes to address the evolving demands of the telecommunications ecosystem.

VISION

The vision of NGMN is to provide impactful industry guidance to achieve innovative, sustainable and affordable mobile telecommunication services to meet the requirements of operators and address the demands and expectations of end users. Key focus areas include Mastering the Route to Disaggregation, Green Future Networks and 6G, while supporting the full implementation of 5G.

MISSION

The mission of NGMN is:

- To evaluate and drive technology evolution towards the three **Strategic Focus Topics:**
  - **Mastering the Route to Disaggregation:**

    Leading in the development of open, disaggregated, virtualised and cloud-native solutions, moving towards agentic AI-based operating models.
  - **Green Future Networks:**

    Developing sustainable and environmentally conscious solutions.
  - **6G:**

    Providing guidance and key requirements for design considerations and network architecture evolution.
- To define precise functional and non-functional requirements for the next generation of mobile networks.
- To provide guidance to equipment developers, standardisation bodies, and collaborative partners, leading to the implementation of a cost-effective network evolution.
- To serve as a platform for information exchange within the industry, addressing urgent concerns, sharing experiences, and learning from technological challenges.
- To identify and eliminate obstacles hindering the successful implementation of appealing mobile services.