13-10-2025

Intelligent Networks: The Rise of Generative AI in Network Management

Grant Agreement No.:	101194278
Work Package:	WP6
Task Item:	T1
Nature of Document:	White Paper
Dissemination Level:	PU
Document ID:	GN5-2-25-11LABC
Authors:	Sonja Filiposka (UKIM); Dimitrios Pantazatos (NTUA); Pavle Vuletić (UoB); Vincent Burkard (FAU); Claudia Torres Perez (i2Cat); Ivana Golub (PCSS)

Abstract

Generative AI is reshaping the landscape of network management, enabling systems that analyse and create configurations, reason on fault resolutions and traffic patterns, and proactively respond to emerging threats. This white paper provides a state-of-the-art review of the use of generative AI through the scope of the FCAPS framework. The aim is to present and highlight current capabilities and research directions, as well as identify the challenges of integrating these technologies into real-world networks.

Co-funded by the European Union

© GÉANT Association on behalf of the GN5-2 project. The research leading to these results has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No. 101194278 (GN5-2).
Co-funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union. The European Union cannot be held responsible for them.

Contents

Figures

Tables

Executive Summary

The recent shift to programmable infrastructure and policy-driven control has created new possibilities for dynamic, distributed, and service-rich network environments. However, this flexibility comes hand in hand with increasing management complexity. Operators are now responsible for a constantly changing ecosystem of devices, services, and constraints. Under these circumstances, traditional network automation and orchestration tools can benefit from the adoption of rapidly emerging Artificial Intelligence (AI)-based tools to help efficiently rise to the new generation of challenges.

This white paper examines how Generative Artificial Intelligence (GenAI) is emerging as a potential supporting solution for the next generation of network management systems. Unlike traditional AI, which predicts or classifies based on input data, GenAI models can generate new, context-aware outputs. This ability enables them to synthesise configurations, simulate faults, and reinterpret telemetry into natural-language explanations.

We provide a survey of the current state of GenAI research and available and emerging tools structured around the ISO FCAPS[1] network management model. We examine how models such as Large Language Models (LLMs), Generative Adversarial Networks (GANs), Variational Autoencoders (VAEs), and Diffusion Models are being applied to networking problems, and we analyse their potential applicability to real-world problems.

The research and tools summary analysis shows that GenAI application in network management looks promising. Research teams report high success rates in configuration generation and fault simulation, with LLM-based systems now reliably translating high-level intent into vendor-specific command syntax. GANs and VAEs are improving the availability and diversity of training data, while diffusion models are emerging as tools for QoS-aware policy synthesis. Hybrid approaches are beginning to show how these models can work together to address complex, multi-step operational tasks.

Across the FCAPS management areas – Fault, Configuration, Accounting, Performance and Security – the most mature and advanced applications of generative AI can be found in configuration and security management. In configuration, LLMs demonstrate strong potential in generating accurate, vendor-specific configurations from natural-language input. In security, generative models create realistic threat scenarios, synthetic attack traces, and anomaly datasets. Fault management is also evolving with generative models that simulate rare faults, generate plausible system logs, and support root cause analysis through co-pilot-style interfaces. In performance management, the use of generative models shows promise in traffic pattern generation, load forecasting, and adaptive retraining, although operational integration remains limited. Accounting management, in contrast, remains the least developed domain, but transferable techniques from other application domains indicate high potential for future development.

It must be noted, however, that GenAI also introduces new challenges and responsibilities. Outputs must be validated and explainable, systems need to be integrated carefully into production environments, and the limits of synthetic data must be understood.

[1] FCAPS is the ISO Telecommunications Management Network model and framework for network management. FCAPS stands for ‘Fault, Configuration, Accounting, Performance and Security’, management categories into which the ISO model defines network management tasks.

1 Introduction

Modern communication infrastructures are no longer monolithic stacks of routers and switches. They span heterogeneous radio, optical and cloud domains, and are orchestrated through layers of intent-driven automation. As a result, administrators must cope with dynamic traffic patterns, rapid service onboarding and ever-shorter control-loop deadlines—challenges that strain rule-based workflows and even classical predictive analytics [1] [2].

These pressures are fuelling demand for more intelligent, adaptive and proactive network-management solutions: platforms that can forecast faults, synthesise remediation policies, and optimise resources before users notice a problem. Early deployments already pair intent engines with machine-learning classifiers, but their effectiveness is often capped by the need for balanced training data and painstaking feature engineering [1].

1.1 Enter Generative AI

Generative artificial intelligence refers to models that not only predict labels from data but also create plausible new samples that follow the same underlying distribution. Where traditional AI answers “what is this packet?”, a generative model can answer “what other packets could exist in this scenario?” and then synthesise them. Prominent families include [3] [4]:

- Large-Language Models (LLMs) that transform text or code sequences.
- Generative Adversarial Networks (GANs) that learn to hallucinate realistic traffic traces or sensor readings.
- Variational Autoencoders (VAEs) that embed network states into smooth latent spaces for anomaly detection or configuration search.
- Diffusion models, which are made robust for constrained optimisation tasks such as wireless contract generation by their iterative denoising.

1.2 Why They Matter for Network Management

Generative models excel when labelled data are scarce, objectives are multi-modal, or the solution space is too ample for exhaustive search [5], which are exactly the pain points of next-generation network operations. Recent studies show diffusion models drafting QoS-aware incentive contracts [6], GANs rebalancing heavily skewed intrusion datasets [7], and LLM co-pilots guiding human operators through complex root-cause analyses [8]. Crucially, pretrained models such as GPT-4, LLaMA-3 or Code-LLM can be fine-tuned on modest domain traces, dramatically lowering entry barriers for operators.

1.3 Aim and Organisation of This Review

This article surveys how generative AI is reshaping network management. We:

- Define the landscape of generative techniques and summarise their theoretical advantages over purely predictive ML.
- Map recent contributions to the classic ISO FCAPS matrix—Fault, Configuration, Accounting, Performance, Security—highlighting where GenAI delivers measurable gains and where gaps remain.
- Identify cross-cutting trends and open challenges, including data-centric AI pipelines, adversarial robustness and ethical considerations.
- Review contemporary commercial and open-source tools in the area of GenAI for network management.
- Distil lessons for practitioners and outline future research directions towards fully autonomous, self-driving networks.

1.4 Scope and Methodology

In the first part of the analysis, we focus on peer-reviewed papers and influential preprints (2018-2025) that explicitly employ generative models for network-management tasks. Sources were selected through research searches using combinations of “generative”, “GAN”, “diffusion”, “LLM”, with “network management”, “orchestration” and FCAPS keywords. Studies were retained if they (i) reported quantitative results on real or emulated networks, (ii) targeted at least one FCAPS function, and (iii) offered code, data or methodological clarity sufficient for reproduction. In the second part, we focus on the analysis of both open-source and commercial tools within the fields of AIOps (Artificial Intelligence for IT Operations) and network performance management.

In the remainder of this white paper, Section 2 offers a concise primer on generative modelling; Sections 3-7 analyse the contributions of GenAI to network management by FCAPS pillar; Section 8 is devoted to an overview of the existing commercial and open-source GenAI tools in the domain of network management; and Section 9 concludes with a summary of the research outlook and key practitioner takeaways.

2 Evolution of Generative AI in Networking

The use of AI in network management has evolved through several key phases [9]. Initially, automation was implemented using rule-based systems that offered fixed logic encoded by network engineers to trigger alarms, apply templates, or activate controls. These systems worked well for predictable scenarios but lacked adaptability to changes. The next wave came with machine learning (ML), where models could learn patterns from data, such as identifying anomalies or predicting link failures. These ML approaches are still very effective in specific tasks, but they are largely reactive and much dependent on labelled data and statically chosen features.

The latest shift in network management approaches is toward generative AI, which enables models not just to classify or predict but to create. Generative AI can synthesise configurations, simulate traffic, draft documentation, and interact in natural language. Hence, this shift marks a fundamental transformation in how automation and decision-making can be approached in network operations.

Unlike earlier systems that operated within predefined boundaries, generative models can propose novel, context-aware solutions tailored to unseen conditions. This allows automation to move beyond static playbooks and into adaptive workflows that can evolve with the network. In decision-making, GenAI enables faster exploration of alternatives, extended “what-if” analysis, and human-friendly interaction through conversational interfaces. It introduces the possibility of co-piloted operations, where AI assists rather than replaces human experts to reduce response times and enhance control-loop effectiveness.

This evolution from rigid rules to reactive learning and on to proactive synthesis is changing the design of network management systems, aligning automation more closely with operator intent, operational risk, and service-level objectives.

2.1 Generative AI Models for Network Management

Generative AI models are mainly designed to produce new outputs that are coherent, realistic, and consistent with learned information. In the context of network management, these outputs may include configuration scripts, synthetic traffic patterns, fault logs, anomaly scenarios, or natural-language summaries [10]. Among the most widely explored families of generative models are Large Language Models (LLMs), Generative Adversarial Networks (GANs), Variational Autoencoders (VAEs), and diffusion models. Each offers unique mechanisms and advantages for addressing different challenges in network operations.

Large Language Models, or LLMs [11], are built on the so-called transformer architecture, which enables them to process long sequences by learning the relationships between tokens using self-attention mechanisms. Trained on massive corpora that include natural language, programming code, and technical documentation, LLMs are very good at contextual reasoning and generating (predicting) coherent sequences of structured or unstructured text. In the networking context, LLMs can be applied to translate operator intent into CLI commands, generate vendor-specific configuration templates, explain alarm outputs in plain language, or assist operators through interactive chat interfaces.

There are several ways to tailor LLMs for specific networking use cases. **Prompt engineering** is the simplest method, involving carefully crafted input queries that guide the model’s behaviour without modifying its internal configuration. Using **fine-tuning**, one can retrain the model using domain-specific examples, including configurations, annotated incident logs, or policy documents. This tactic lets the model specialise in tasks relevant to a particular network environment. A third strategy is **Retrieval-Augmented Generation** (RAG), which integrates the LLM with an external knowledge source, such as a network documentation repository. At inference time, the system retrieves relevant context (for example, a recent ticket or device specification) and feeds it to the LLM, allowing for more accurate, specific and up-to-date responses.

Generative Adversarial Networks [12] operate through an adversarial framework involving two neural networks: a generator, which produces synthetic data, and a discriminator, which attempts to distinguish generated data from real examples. As both networks learn simultaneously, the generator improves its ability to produce high-fidelity samples. In network management, GANs have been widely adopted for generating synthetic traffic data, simulating rare fault conditions, and producing labelled intrusion examples for training security systems. These capabilities are especially valuable in scenarios where real-world data is limited or imbalanced, such as training datasets for anomaly detection or failure prediction. Although GANs can achieve impressive realism, they require careful calibration to avoid problems such as mode collapse, where the generator learns to produce only a narrow subset of outputs rather than covering the whole variety of the training data.

Variational Autoencoders [13] are probabilistic models that learn a compact, latent-space representation of the input data. An encoder maps inputs to this space, while a decoder reconstructs them, allowing for controlled sampling and interpolation between valid states. VAEs have found application in modelling the distribution of valid configurations, detecting anomalies via reconstruction error, and generating intermediate network states for smooth policy transitions. Their structure encourages interpretability and stability, making them especially useful in environments where explainability and safety are critical.

Diffusion models [14] are a newer class of generative models. They function by gradually adding noise to input data through multiple steps, then learning to reverse this process to generate structured outputs. This iterative denoising mechanism leads to high-quality results and allows for precise control over the generative process. In network operations, diffusion models show promise in synthesising configurations that meet specific performance or QoS (Quality of Service) constraints and generating policies under operational constraints. While they offer strong robustness and control, diffusion models are more computationally intensive than other generative approaches.

An increasingly important direction in the application of generative AI is the combination of these model types into **hybrid pipelines**. Examples of these hybrid approaches include:

- **LLM + GAN**, where an LLM generates fault narratives or operator instructions based on synthetic logs produced by a GAN.
- **LLM + VAE**, where a VAE maps the valid configuration space and an LLM converts operator intent into traversable latent vectors.
- **LLM + RAG + diffusion**, where an LLM uses retrieval to ground configuration goals, which are then passed to a diffusion model to synthesise a matching configuration under QoS constraints.

2.2 Benefits and Challenges

The application of generative AI in network management can introduce several key benefits across technical, operational, and human-centric dimensions, offering capabilities that go well beyond the limitations of rule-based or purely predictive systems.

One of the most immediate advantages is **scalability**. Once trained or adapted, generative models can produce valid device configurations, synthetic data samples and diagnostic suggestions fast, at scale and on demand. This eliminates many of the repetitive, manual, time-consuming tasks currently implemented by operations teams, particularly during large-scale deployments, migrations or testing.

Generative models can offer greater **adaptability**, especially in scenarios with little labelled data, high system dynamics or anomalies. In these cases, generative AI can respond flexibly to the changing conditions even in the face of system states that it has never seen before.

Another key benefit lies in **interactivity**. LLMs allow operators to engage with network systems using natural language. This can dramatically lower the barrier to automation, especially for teams that lack expertise in scripting and development skills. This interaction also improves **transparency**, allowing the operators to request explanations instead of blindly trusting a black-box execution. LLMs can also simulate changes or query compliance status in conversational form, enabling the practical development of AI co-pilot systems that support human decision-making.

Generative models further support **faster prototyping and simulation**. Models such as GANs or diffusion architectures can generate synthetic traffic traces, emulate rare network events, or simulate operating scenarios under policy constraints. This is especially useful for stress-testing, evaluating SLAs under future load conditions, or training models without risking live infrastructure. Because generative models can create diverse and controllable scenarios, they are powerful tools for “what-if” analysis in performance, fault or capacity planning.

However, as is often the case, generative AI can be a double-edged sword – despite the advantages it offers in network management, there are also important challenges and risks associated with its use.

One key concern is the **validity and safety** of generated outputs. Even small inaccuracies in automatically generated configurations or security responses can lead to outages or compliance violations. LLMs, for example, may hallucinate seemingly valid but, unfortunately, incorrect commands if not carefully prompted. Similarly, GANs might produce unrealistic traffic patterns if trained on biased or insufficient data. These issues showcase the need for robust post-generation validation mechanisms, safety constraints, and, in many cases, human-in-the-loop review.

Another significant challenge is **explainability**. Many generative models operate as black boxes, providing limited insight into why a specific output was generated. This creates a considerable problem for production environments where accountability, traceability, and auditability are essential. Operators would be reluctant to trust or adopt generative outputs that cannot be explained, especially in domains such as security or policy enforcement.

Integration complexity is additionally challenging when it comes to generative AI models and their need to interact with other systems, real-time data sources, orchestration pipelines, and monitoring platforms. Successful implementation requires technical compatibility and operational alignment to ensure that the AI-generated outputs can be consumed, validated, and acted upon.

Concerns exist regarding the **use of synthetic data**. Poorly calibrated generative pipelines can introduce bias, artefacts, or unrealistic conditions into model training. This can degrade the performance of the rest of the systems and potentially increase inaccuracies or unfairness instead of fixing the issues. Ensuring realistic representation in the generated data remains an open research area.

The following sections examine how generative AI capabilities are applied within the FCAPS framework domains (Fault, Configuration, Accounting, Performance, and Security management). In each domain, we highlight which model families are most prevalent, what tasks they support, and how their integration enhances the intelligence, autonomy and resilience of network management systems. While generative AI introduces tremendous benefits to network management, such as automation, simulation, reasoning and interaction, it also brings risks and challenges that will be explored in more detail across the FCAPS domains.

3 Fault Management

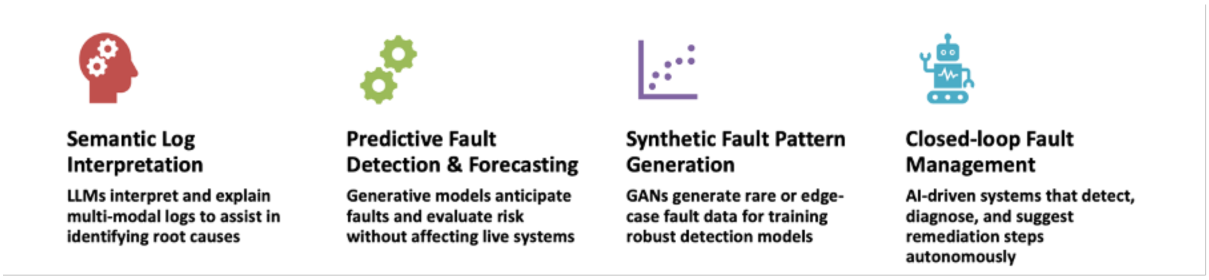
Fault management is a critical part of network operations, aiming to detect, diagnose, and resolve problems with minimal impact on services. With networks generating increasingly complex and multi-modal data, the ability to detect, diagnose and respond to faults quickly and accurately is essential to maintaining service reliability. Traditional fault management systems rely on rule-based monitoring, manual log analysis, and static alerting systems, and hence struggle to cope with the volume, velocity, and variety of modern network data.

Recent advances in Generative AI can be used to synthesise human-readable explanations, generate training data for rare fault conditions, and provide semantic interpretations of complex logs and multi-modal inputs. As a result, Generative AI is beginning to transform fault management from reactive troubleshooting into proactive, context-aware, and even predictive network resilience.

3.1 Opportunities for Generative AI

Generative AI offers several promising avenues to enhance fault management in network operations. One key opportunity lies in the semantic interpretation of log data, where LLMs can extract meaning from unstructured inputs, helping engineers identify patterns and understand faults more quickly. Unlike static keyword-based approaches, LLMs can handle ambiguity, correlate symptoms, and summarise root causes in natural language.

Another emerging area is the generation of synthetic fault data. GANs are being used to simulate rare or complex failure scenarios, helping to address the long-standing issue of imbalanced datasets in training fault detection systems. This enables more robust models that generalise better to edge cases.



Semantic Log Interpretation

LLMs interpret and explain multi-modal logs to assist in identifying root causes

Predictive Fault Detection & Forecasting

Generative models anticipate faults and evaluate risk without affecting live systems

Synthetic Fault Pattern Generation

GANs generate rare or edge-case fault data for training robust detection models

Closed-loop Fault Management

AI-driven systems that detect, diagnose, and suggest remediation steps autonomously

Figure 3.1: Opportunities for GenAI in fault management

In more advanced implementations, generative AI supports proactive and predictive fault handling. By combining generative models with digital twins, it is possible to simulate future failure scenarios, forecast network health, and test recovery strategies—all without affecting live infrastructure. This capability is especially relevant in large, dynamic environments where real-time experimentation is impractical.

Finally, generative AI opens the door to closed-loop fault management systems, where detection, diagnosis, and remediation are handled autonomously with human oversight. These systems promise not only faster resolution but also reduced operational overhead and more consistent incident handling.

3.2 Review of GenAI Techniques and Models

The reviewed work presented in Table 3.1 below showcases a range of generative AI techniques being applied to fault management in network operations. Across these studies, the use of LLMs, GANs, and variational autoencoders (VAEs) forms the core of current innovations.

Reference	Approach	Experiment Setup	Success Level	Key Findings
LLM-Assisted End-to-End Network Health Management Based on Multi-Scale Semanticization [15]	Multi-Scale Semanticized Anomaly Detection with LLMs	NS3 simulations of 7 scenarios of heterogeneous networks, including base stations and vehicle networks, with six fault categories + open dataset for intrusion detection – evaluated against state-of-the-art time-series anomaly detection schemes (CL-MPPCA, SR-CNN, ANOMALYBERT, MULA, Swinlstm, TranAD).	Achieved 97.09% anomaly detection accuracy, outperforming other schemes by at least 24% and 89.42% fault classification accuracy & outperforming traditional engines by 22%.	Integrates semantic rule trees and LLMs for anomaly lifecycle management from detection to mitigation.
Integrating Generative AI with Network Digital Twins for Enhanced Network Operations [16]	Integration of GANs/VAEs with digital twins	A scalable model of a 6G network environment is simulated to test under various predefined scenarios, including peak usage times, network failures, and security breaches.	Latency, bandwidth, error rate and detection time are improved. Enhanced predictive accuracy to 95%. Network failure recovery time is 3x faster, and security detection time is 4x faster.	Combining GenAI with digital twins enhances situational awareness and proactive fault handling.
When Digital Twin Meets Generative AI: Intelligent Closed-Loop Network Management [17]	GenAI-empowered Digital Twin (GDT) architecture	A case study on data-model-driven network management for mobile networks focusing on QoE. A multicast short video-streaming scenario with two APs, multiple multicast groups and one GDT.	It can always maintain the highest QoE compared to the heuristic and DRL approach.	A GDT network architecture can achieve intelligent external and internal closed-loop network management.
Adapting Network Information into Semantics for Generalizable and Plug-and-Play Multi-Scenario Network Diagnosis [18]	NetSemantic framework utilising LLMs with semanticisation and symbolisation of multimodal network data	Network fault dataset with six major network failures collected from a digital twin experiment with an embedded semi-physical platform and network simulator NS-3 using 9 to 20 nodes. Results compared with several popular ML/DL-based baselines.	Demonstrates superior performance on several key indices, achieving a 5-10% improvement over baseline. The recognition rate of anomalous samples reaches up to 98%. Fault diagnosis reaches up to 96.1% accuracy with GPT4.	Introduces a plug-and-play data-independent framework that transforms network information into unified textual representations, enabling zero-shot fault diagnosis across various network environments.

Table 3.1: Comparative summary of recent papers on the topic of GenAI for fault management

In *Plug-and-Play Multi-Scenario Network Diagnosis* [18], the authors apply transformer-based LLMs to interpret unstructured log and telemetry data. These models are used not only for fault detection but also for explainability, producing interpretable summaries and root cause insights from complex data streams. For example, the latter paper introduces a framework that transforms network data into semantically enriched textual inputs, enabling zero-shot generalisation across diagnostic scenarios.

In parallel, in *Integrating Generative AI with Network Digital Twins for Enhanced Network Operations* [16], the authors leverage GANs and VAEs to generate synthetic fault scenarios. These models help mitigate the problem of imbalanced fault datasets, improving the robustness of classifiers in low-frequency failure conditions. In environments with rare or insufficient fault samples, the generated synthetic data can expand training datasets, enhancing resilience to edge-case failures. However, the realism of synthetic data remains an open concern, particularly when models are deployed in high-stakes operational environments.

Several papers integrate GenAI with simulation environments, building digital twin systems. *LLM-Assisted End-to-End Network Health Management Based on Multi-Scale Semanticisation* [15] introduces a multi-scale semanticised anomaly detection model that combines LLMs with hierarchical attention mechanisms. Using attention with semantic rule trees, the authors strive to detect and forecast anomalies in simulated heterogeneous networks. In *When Digital Twin Meets Generative AI: Intelligent Closed-Loop Network Management* [17], the research goes one step further by proposing a conceptual framework for intelligent closed-loop autonomous fault handling, though it is still in the early validation stages.

Preprocessing is a consistent and crucial step across all implementations. The developed pipelines include log structuring, feature extraction, and symbolic representation to prepare the data for the models. This additional layer improves the quality and alignment of generative outputs with operational expectations, ensuring that the models can produce contextually relevant and actionable insights in real-world diagnostics.

3.3 Strengths and Open Challenges

Across the reviewed studies, generative AI demonstrates several clear strengths in fault management. Most notably, it enhances explainability, especially when using transformer-based LLMs that can interpret network logs and translate technical anomalies into understandable narratives. This ability supports root cause analysis and reduces the cognitive load on operations teams.

Another strength is proactive fault prediction. Generative models integrated with digital twins can be used not just for detecting faults after they occur but also for simulating “what-ifs” and future scenarios and forecasting network problems. This capability enables proactive intervention, efficient maintenance planning, and reduced downtime.

The use of GANs and VAEs also contributes to improved robustness of the detection models. By generating synthetic fault conditions, these models help overcome dataset imbalances and prepare classifiers for rare or novel incidents, traditionally an area where non-AI-based systems often underperform.

However, while the research results are promising, e.g., 97% detection accuracy in one study, practical deployment still faces challenges. Issues such as latency and scalability must be considered for real-time fault detection systems that must operate under strict performance constraints. Most current GenAI prototypes have not been tested in such settings. Another key concern is the realism and operational reliability of synthetic data. While generated faults can expand training sets, their accuracy and relevance in live environments need closer validation.

Integration complexity is another barrier. Most reviewed systems are developed in simulated or controlled testbeds, and deployment in multi-vendor, real-world networks remains limited. This includes aligning AI-generated insights with human oversight, organisational policies, and fault-handling workflows. Finally, generalisation across domains is still evolving. While semantic models show promise in adapting to new inputs, further work is needed to ensure consistent performance across technologies, topologies, and data formats.

4 Configuration Management

With the evolution of increasingly dynamic and complex networks, engineers are shifting to focus on automation and orchestration solutions, as traditional manual configuration has become unmanageable. Recent developments in generative AI offer the possibility to further streamline the process of network configuration and verification. By learning from existing examples and interpreting high-level input such as natural language or policy intent, generative models have the potential to support network engineers in producing accurate, consistent, and standards-aligned configurations.

4.1 Opportunities for Generative AI

The most prominent use case for integrating generative AI into network configuration management is to translate high-level intent into device-level configurations, following the principles of intent-based networking (IBN). Instead of focusing on how to automate typical network engineering tasks using platform-specific configuration code, generative AI goes one step beyond and aims to interpret what the user wants and produce a configuration with valid syntax and complex semantic meaning.

Several recent research studies demonstrate that one can use generative AI to generate device configurations and assist with policy management successfully. One of the most widely explored use cases is translating natural-language intent into vendor-specific configuration files. When provided structured inputs or standards-based templates, models can generate valid vendor-specific configurations from plain-text instructions. These efforts highlight the feasibility of generative AI as an intelligent intermediary between operator goals and device-level implementation. These efforts lead towards true zero-touch provisioning, where AI can handle basic connectivity or routing setup with minimal human input. These use cases not only speed up deployment but also reduce configuration errors and make advanced networking features more accessible to less experienced operators.

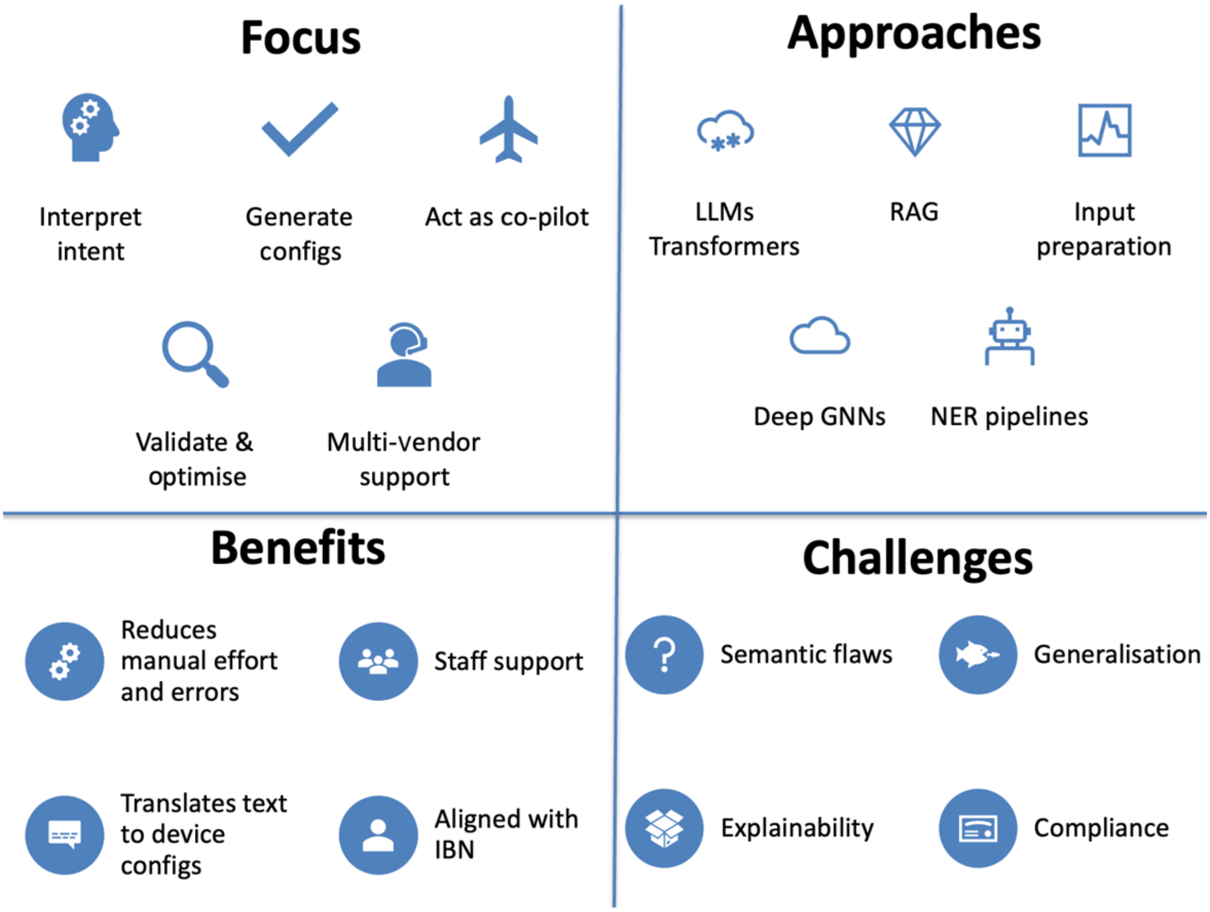


Figure 4.1: Potential uses of GenAI in configuration management

Another opportunity lies in creating AI-powered configuration assistants, or "co-pilots." Rather than replacing human decision-making, these AI tools are designed to guide, support, and explain. They offer suggestions, explain configuration logic, retrieve relevant documentation, and adapt to specific vendor environments. The interaction is often envisioned as conversational, allowing engineers to iteratively improve results and clarify their intent. This model of support is especially valuable for less experienced operators or for teams managing multi-vendor environments where syntax and semantics can vary widely. The collaborative nature of the approach preserves human oversight while significantly reducing cognitive load and time-to-solution.

Generative AI is also used for automated validation and optimisation, moving beyond text generation into intelligent decision support. Rather than simply generating configurations, AI systems can validate them against policies or goals, flag potential issues, or suggest improvements. In this context, it is essential to develop benchmarking that will ensure that AI systems can deliver not only syntactically valid outputs but also ones that align with intent, compliance, and operational safety.

Finally, AI models are increasingly used to bridge the gap between multi-vendor and hybrid environments. Techniques for abstracting intent and producing configurations for different platforms by incorporating vendor-aware preprocessing and fine-tuned pipelines can enable unified management across different infrastructures. The power of these approaches lies in tailoring outputs while maintaining a unified interface. This ability to generate configurations across heterogeneous environments is crucial for scaling automation in complex networks.

4.2 Review of GenAI Techniques and Models

Researchers are experimenting with various generative approaches, including GANs, autoencoders, and most prominently, LLMs. GANs, while powerful in modelling patterns, seem less suited to structured tasks such as configuration generation, where precision, interpretability and user interaction are critical. Autoencoders, similarly, are helpful for anomaly detection or feature extraction, but are not designed for intent-to-configuration workflows.

By contrast, LLMs, particularly if built on transformer architecture, have rapidly emerged as the leading approach for intent-driven network configuration. Specifically, the transformer model processes sequences of information by attending to all parts of the input simultaneously, rather than step by step. This makes it highly effective at understanding context and relationships across a configuration instruction or template. If trained on a large corpus of network documentation and configuration examples, LLMs have the potential to generate complete, coherent configuration snippets based on natural language prompts or abstract policy descriptions. Moreover, their interactive nature allows them to support human-in-the-loop workflows, making them especially suitable for real-world operational use.

Effective utilisation of LLMs often involves prompt engineering, designing prompts or input instructions in ways that optimise model output. Prompt engineering can significantly influence the accuracy, coherence, and relevance of the generated outputs, making it an essential skill for leveraging LLMs in operational scenarios. By carefully crafting prompts, practitioners ensure that LLM-generated configurations align closely with user intent and operational requirements.

Reference	Approach	Experiment Setup	Success Level	Key Findings
Generative AI for Low-Level NETCONF Configuration in Network Management Based on YANG Models [19]	LLMs with YANG Models integration	Network with 7 hosts, 2 switches and 2 routers. Tests based on 12 relatively simple configuration scenarios.	Best model (GPT-4o with pipeline) achieved 57% valid tests. No scenario is always correct; more complex scenarios are always wrong. Problems include missing or misused XML tags, incorrect tree hierarchy, and syntax errors.	LLMs show promise for NETCONF configuration generation, but are not yet mature enough for complex industrial applications.
A Novel LLM Architecture for Intelligent System Configuration [20]	LLM-based intelligent chatbot architecture with RAG and function calling	3 domain-specific tasks with 30 sample conversations each: static NAT conf, fake server conf, assistant API conf	The success rate is 100% for NAT and 50% for server conf. Assistant API conf requires postprocessing and future work.	LLM architectures show potential for learning assistance. IBN with natural language is defined as a future step.
NetConfEval: Can LLMs Facilitate Network Configuration? [21]	LLM, GPT-4	Network size varies from 33 routers to 2 for routing (OSPF, RIP, BGP) using emulation. Input: max 3200 network high-level requirements. Tasks: generating formal specification, generating API calls, developing routing algorithms, and generating low-level configuration (max 10 lines)	GPT4 achieves almost 100% accuracy, but all models are bad at complex conflict detection. Native-function calling does not perform as well. Models cannot directly calculate routing paths using the shortest-path policy for a network of 10 devices. GPT4 can accurately generate routing code when given feedback. Error-free low-level configurations in only 1 of 3 runs. With RAG, it can reach 100% accuracy, but not for OSPF.	LLMs show potential in facilitating various network configuration tasks, but complex tasks need to be split into smaller subtasks, task-specific verifiers are needed, and humans must still be in the loop.
Large Language Models for Zero Touch Network Configuration Management [22]	Local LLM (Zephyr-7b) with verification and orchestration modules	Network: a partial mesh topology with four Cisco routers and two hosts. Dataset with 90 resources. Configuration intents divided into four types (conf properties, routing, ACL, tunnel conf).	LLM-NetCFG correctly classified and generated the correct configuration for 92.2% of requirements. Successful handling of non-complex intents in 5-7 minutes.	LLM-NetCFG demonstrates the capability of generating and verifying network configurations based on natural language intents. LLMs can be used to design ZSM self-configuration agents.
Towards Intent-based Network Management for the 6G System adopting Multimodal Generative AI [23]	LLMs with industry-ready standard templates: TM Forum and GSMA Open Gateway	Network: Patras5G testbed [24] Input: 3 classes (massive IoT, ultra-low latency comm, enhanced mobile broadband) with five prompts each.	Proof of concept works.	Proof of concept translation of high-level intent into network configurations (TMF Service Order provided to OpenSlice OSS that uses the Operate API to push the configuration).
Mobile Network Configuration Recommendation Using Deep Generative Graph Neural Network [25]	Deep Generative Graph Neural Network (GNN) with Siamese architecture compared to Graph Auto-Encoder (GAE)	Datasets from 2 networks created using collected information from multiple eNBs and gNBs, combining LTE and NR attributes.	GAE excels in unsupervised learning from non-configuration data (99.1% acc), S-GNN is more accurate in supervised scenarios (92.1% acc), but may face challenges in adapting network configuration changes.	Deep Generative GNNs can recommend network configuration parameters and detect misconfigurations.

Reference	Approach	Experiment Setup	Success Level	Key Findings
Making Network Configuration Human Friendly [26]	LLM, GPT-4	Network: 7 P4 switches with two hosts Task: MPLS routing	Capable of handling up to 40 requirements per prompt accurately; for higher numbers, it becomes unstable. In some experiments, generated configs are semantically correct, but contain some syntax errors. No complete accuracy evaluation provided.	NETBUDDY can generate working P4 and BGP configurations from high-level requirements.
LLMNDC: A Novel Approach for Network Device Configuration based on Fine-tuned Large Language Models [27]	Fine-tuned LLMs	Network: Huawei and Cisco devices Task: query communication knowledge and generate configuration suggestions.	Fine-tuned model effectively generates accurate configurations for various network devices, demonstrating improved performance over baseline models.	Fine-tuned LLMs can provide configuration recommendations for network devices.
Automation of Network Configuration Generation Using Large Language Models [28]	Extract keywords using NER, send to LLM, map to key-values and standardise the orch API.	Input: different tariff plans of 27+ operators from which five information elements are identified. Output: API call to network orchestrator that forms the relevant configuration.	A fine-tuned GPT-3.5 model with an input filter attained a perfect F1 score of 1, compared to 0.962 without it. The Llama2-7B model also showed significant improvements with the input filter.	LLM-based pipeline converts input tariff plans and other additional parameters to related configuration and provisions complex networks.
Can LLMs Understand Computer Networks? Towards a Virtual System Administrator [29]	Evaluation framework testing six LLMs (proprietary and open source) on tasks: topology analysis, IP recognition, and path computation.	3 networks: 2 devices, 3 routers and 5 subnets; 12 nodes and 15 subnets. 13 different tasks divided into 4 groups: topology, drawing, addresses, and paths.	Zero-shot Bing achieved a mean 79% accuracy; open-source models showed lower performance (Llama 2 is worst with 37% mean accuracy), especially on complex networks.	LLMs can effectively handle basic network tasks in simple topologies, but their performance declines with increased complexity. Prompt engineering can enhance accuracy.
What do LLMs need to Synthesize Correct Router Configurations? [30]	Verified Prompt Programming—combining GPT-4 with automated verifiers and localised feedback mechanisms.	Task 1: Translate Cisco to Juniper configuration, including BGP, OSPF, prefix lists, and route maps. Task 2: Generate router configs for a given network topology based on local policies.	Leverage of 10x for Juniper translation, and 6x for implementing the no-transit policy. Stand-alone works badly, making elementary errors that can bring networks down.	LLMs alone are insufficient for accurate configuration synthesis; integrating verifiers and providing modular, localised feedback substantially improves outcomes.
S-Witch: Switch Configuration Assistant with LLM and Prompt Engineering [31]	Combines an LLM with a digital network twin (in GNS3) to generate and verify switch configurations through prompt engineering.	Network: 3 routers, 2 switches, 5 hosts Input: high-level requirements in natural language.	IP allocation was incorrect for some of the devices. Blocking traffic with ACL was successful. Setting up VLAN was with mixed success: trunk mode works only if explicitly specified.	Integrating LLMs with network digital twins and prompt engineering can automate configuration tasks, but additional methods, such as RAG, must be used to overcome limitations.

Table 4.1: Comparative summary of recent papers on the topic of GenAI for configuration management

The majority of recent papers on using generative AI for network configuration focus on LLMs. For example, the study *Generative AI for Low-Level NETCONF Configuration Management* combines LLMs with structured YANG models and natural language to generate NETCONF-compliant XML configurations. Another example is *Towards Intent-Based Network Management for the 6G Era* [23], which uses LLMs in combination with standardised TM Forum and ETSI templates to transform high-level policy intent into complete configurations.

Newly developed co-pilot systems, such as in *A Novel LLM Architecture for Intelligent System Configuration* [20] and *Can LLMs Understand Computer Networks?* [29], implement retrieval-augmented generation to assist users. These systems first fetch relevant examples or prior configurations from a knowledge base and then use the LLM to generate context-aware suggestions. This provides for more reliable solutions and a clearer audit trail. Another variation is presented in *Large Language Models for Zero Touch Network Configuration Management*, which leverages a locally deployed LLM that can handle non-complex configuration tasks using on-premise orchestration. Together, these approaches demonstrate the flexibility of generative AI across centralised and distributed deployment models.

Several approaches go beyond general-purpose models by using fine-tuned LLMs trained on domain-specific data. For instance, *LLMNDC* fine-tunes LLMs to generate network device configurations with improved accuracy, and the study *Automation of Network Configuration Generation Using Large Language Models* applied a named-entity recognition (NER) preprocessing step to extract relevant information from user inputs before mapping to configurations using an LLM. Another notable example is *S-Witch*, which combines an LLM with a network digital twin and rule-based validation system to generate executable and verified CLI commands for commercial switches.

In some cases, prompt design and post-processing are key to success. *What Do LLMs Need to Synthesize Correct Router Configurations?* presents a Verified Prompt Programming approach, which combines high-precision natural-language prompts with post-generation validation. *Making Network Configuration Human-Friendly* translates high-level policies into valid P4 and BGP configurations through a proof-of-concept LLM pipeline. In fact, most systems employ some form of preprocessing before input reaches the model. This may include translating user input into structured prompts, retrieving similar examples, validating syntax constraints, or inserting vendor context. These steps are critical to model performance, ensuring outputs are aligned with operational goals and syntax standards.

Evaluation methods vary but often include performance benchmarking across models (NetConfEval, *Can LLMs Understand Computer Networks?*), task-specific experiments using real-world configuration sets (LLMNDC, S-Witch), or cross-vendor translation tests [30]. These evaluations highlight the progress made and the limitations that still need to be addressed. Regarding success rates, results show moderate to high accuracy depending on task complexity and preprocessing quality. While performance is promising for straightforward configuration tasks, success levels decline in cases involving complex logic, multi-vendor abstraction, or ambiguous policy intent.

4.3 Strengths and Open Challenges

Generative AI offers a powerful new paradigm for simplifying and accelerating network configuration workflows. Across the reviewed studies, a consistent strength is the ability of models, especially fine-tuned LLMs and retrieval-augmented systems, to translate high-level, human-friendly inputs into actionable configurations for real-world network devices. In use cases such as NETCONF, P4/BGP, and 5G provisioning, generative systems reduce the need for manual scripting and allow engineers to express intent in natural language, with the AI handling syntax, formatting, and vendor-specific conventions.

Across nearly all of the reviewed studies, there is a consistent emphasis on preprocessing and input structuring before LLMs are invoked. Whether parsing user intent into formal templates, selecting vendor context, retrieving past configuration examples, or encoding relevant policies, this preprocessing step dramatically improves AI-generated outputs' quality, relevance, and safety. It also addresses a key limitation of generic LLMs—their lack of awareness of operational constraints or enterprise-specific standards. This layered approach, where preprocessing enriches the input and postprocessing validates the output, is emerging as a best practice in real-world AI-in-the-loop deployments.

A consistent and important design principle across all studies is the emphasis on keeping humans in the loop as a deliberate architectural choice. The reviewed systems incorporate co-pilot models that support, rather than replace, engineers. Most tools rely on interactive prompts, output previews, and configuration suggestions rather than automatic enforcement. Several papers also underline the limitations of current models, such as occasional hallucinations, vendor-specific misunderstandings or ambiguous intent interpretation, which further necessitate human oversight. In more advanced systems, AI-generated configurations undergo additional verification steps or policy checks before deployment, reinforcing the engineer's role as the final authority.

The reviewed work also shows that generative AI can adapt to diverse environments, from programmable switches and 5G service platforms to multi-vendor routing backbones. This versatility is enabled by fine-tuning, task-specific preprocessing, and modular design, allowing generative systems to understand context, retrieve relevant knowledge, and tailor their outputs accordingly.

Despite these promising capabilities, several challenges and limitations persist. A key concern, echoed in papers like *NetConfEval*, *Can LLMs Understand Computer Networks?*, and *What Do LLMs Need to Synthesise Correct Router Configurations?*, is that LLMs can often produce syntactically correct but semantically flawed configurations. The accuracy and reliability of LLMs still vary significantly based on the complexity and ambiguity of the input, especially in multi-domain or mission-critical environments. Many models struggle with nuanced requirements, edge cases, or uncommon syntax variations. Without adequate prompting, context, or validation, models may hallucinate commands, misinterpret policy intent, or mix configuration styles from different vendors.

Another recurring issue is generalisation and reliability. Many models perform well in narrow or controlled benchmarks but struggle in real-world scenarios where edge cases and inconsistent documentation are common. Papers like *Making Network Configuration Human-Friendly* and *LLMNDC* show that fine-tuning improves accuracy, but also emphasise that training data quality and diversity remain a great problem in this area.

Moreover, compliance and auditability are areas where current tools fall short. While some systems incorporate policy-aware validation, there is no unified framework for aligning AI-generated outputs with enterprise governance, change management practices, or external compliance standards. This gap raises concerns about accountability, especially in regulated or safety-critical environments.

Finally, explainability and trust are pressing challenges, as many LLMs still function as black boxes. Understanding why a model proposed a specific configuration or verifying that it fully aligns with the original intent remains complex problems, especially for users without deep AI expertise.

In summary, generative AI systems for network configuration are evolving quickly, with clear strengths in flexibility, time savings, and usability. However, they are not yet production-ready for fully autonomous deployment in most environments. As these tools move closer to deployment, questions around accountability, explainability, and safe rollout of configurations at scale remain largely unresolved. Addressing these gaps will be critical to ensure that generative AI becomes not only a helpful assistant but also a trusted and integral part of network automation workflows.

5 Accounting Management

Network accounting refers to the monitoring, measuring, and analysing of resource usage across a network, typically for billing, cost optimisation, auditing, and service-level enforcement. It is central in usage-based pricing models, dynamic charging schemes, and performance-linked service guarantees. Traditional network accounting systems rely on predefined rules and structured records, such as call detail records (CDRs) or IP flow data, and are often limited in their ability to adapt to evolving service models and user behaviours.

Despite the growing interest in applying Generative AI across network operations, current academic research reveals a notable gap in its application where network accounting management, specifically in usage-based billing, dynamic charging models, and revenue assurance are concerned.

To date, no peer-reviewed studies that explicitly focus on using GenAI for accounting tasks in a networking context have been identified. However, related work in financial accounting [32], automated billing systems, and predictive analytics for revenue forecasting [33] demonstrates the potential of LLMs, VAEs, and GANs for generating synthetic transaction data, detecting anomalies, and producing explainable summaries.

Although these techniques are primarily developed for financial applications in other domains, they could be adapted to improve the transparency, scalability, and automation in network accounting systems [34] [35]. Future research must validate their effectiveness in network environments where data heterogeneity, regulatory compliance, and real-time constraints introduce unique challenges.

6 Performance Management

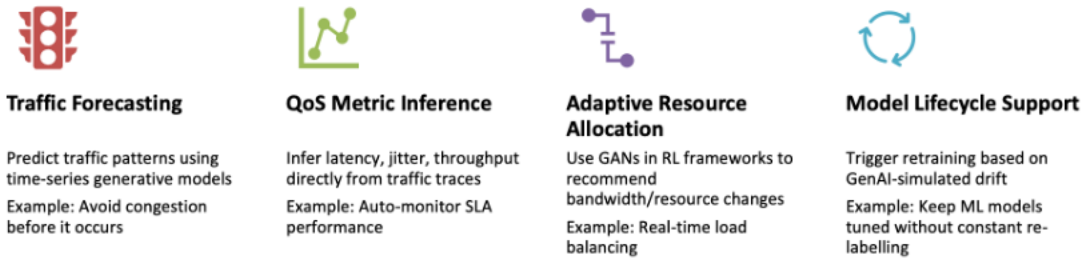
Performance management in networking systems focuses on ensuring the availability, responsiveness and efficiency of network resources. It involves monitoring and optimising key performance indicators such as throughput, latency, jitter, and reliability. These metrics are critical for meeting service-level agreements and user expectations.

Traditional tools, while effective in stable environments, struggle to adapt to the increasing complexity of traffic patterns, dynamic service topologies, and multi-access network layers. Recent advances in Generative AI offer new mechanisms for addressing these challenges. By learning latent performance structures from traffic data, simulating network behaviours under diverse loads, and supporting adaptive decision-making, generative models are beginning to reshape how performance is inferred, predicted and optimised across modern networks.

6.1 Opportunities for Generative AI

Generative AI opens new pathways for more adaptive, efficient, and predictive network performance management. A key opportunity lies in forecasting traffic and performance patterns using generative time-series models. By learning from historical data, these models can anticipate spikes, congestion, or degradation, allowing for proactive adjustments before service levels are impacted.

Another emerging area is the generation of performance visualisations and summaries using LLMs. When integrated with telemetry data, LLMs can produce human-readable performance dashboards, trend analyses, and anomaly reports, reducing interpretation overhead and enhancing visibility in complex environments.



Traffic Forecasting

Predict traffic patterns using time-series generative models

Example: Avoid congestion before it occurs

QoS Metric Inference

Infer latency, jitter, throughput directly from traffic traces

Example: Auto-monitor SLA performance

Adaptive Resource Allocation

Use GANs in RL frameworks to recommend bandwidth/resource changes

Example: Real-time load balancing

Model Lifecycle Support

Trigger retraining based on GenAI-simulated drift

Example: Keep ML models tuned without constant re-labelling

Figure 6.1: Potential uses of GenAI in network performance management

Generative models also play a role in adaptive resource management. In domains such as network slicing and optical networks, GAN-enhanced reinforcement learning systems can simulate varied traffic conditions and help optimise decisions under constraints such as bandwidth limits, routing rules, or SLA targets. This leads to more stable resource allocation and better handling of dynamic usage patterns.

Finally, generative AI provides tools for automated retraining and performance maintenance in ML-based systems. By simulating performance drift or injecting synthetic workloads, generative models can help identify when retraining is necessary and reduce the operational cost of keeping predictive systems up to date.

This way, GenAI goes beyond monitoring performance to actively shape traffic via simulation, prediction, explanation, and control.

6.2 Review of GenAI Techniques and Models

There are still relatively few published research papers focusing specifically on GenAI's application to network performance management. The existing literature offers promising starting points, but this remains an emerging research area with considerable room for future exploration. The selected papers demonstrate a diverse application of generative AI techniques across network performance management scenarios, ranging from QoS inference to adaptive retraining and optical network optimisation. Together, they showcase how generative models can improve performance prediction, resource control, and system robustness.

Reference	Approach	Experiment Setup	Success Level	Key Findings
Deep Generative Model and Its Applications in Efficient Wireless Network Management: A Tutorial and Case Study [36]	Using Deep Generation Models for improving the efficiency of wireless network management	Case study: simulation of mobile user behaviour and Artificial-Intelligence Generated Content (AIGC) incentive contracts under QoS constraints.	The approach achieves stable contract generation, enabling clients to always obtain positive utilities while maintaining high generation quality.	Deep generative models, particularly diffusion models, can be applied to optimise QoS-constrained wireless service models through incentive design.
Deep-Q: Traffic-driven QoS Inference using Deep Generative Network [37]	Deep generative network for direct inference of QoS metrics from raw traffic traces (VAE + LSTM)	Network testbeds of two different topologies: data centre network with six switches and overlay an IP network with 8 servers.	Deep-Q achieves on average 28% lower inference errors than VAE. Keeps a stable average inference error below 15% in all cases.	To achieve a high inference accuracy, combine VAE's advantage of stable training performance and high modelling accuracy of GANs.
Generative-AI for AI/ML Model Adaptive Retraining in Beyond 5G Networks [38]	Generative AI-based (VAEs + GANs) framework to trigger ML model retraining based on drift detection and relevance analysis	Evaluated for two use-cases: QoS prediction over the O-RAN software community platform, and network slicing using a real-time dataset.	Reduced model retraining overhead while maintaining >95% performance accuracy.	GenAI supports adaptive ML pipeline management by simulating performance degradation patterns for proactive retraining.
GAN-powered Deep Distributional Reinforcement Learning for Resource Management in Network Slicing [39]	GAN-enhanced deep distributional Q network [40] for managing resource allocation across network slices	Simulated multi-slice network with varying traffic and SLA constraints. A RAN scenario with 3 service types (VoLTE, video, and URLLC) and 3 slices in each BS. 100 registered subscribers.	Duelling GAN-DDQN performs significantly better, improving system utility, bandwidth allocation, and offering a Service Satisfaction Ratio of 1.	It is non-trivial to optimise multiple conflicting objectives, even when using cutting-edge algorithms.
OpticGAI: Generative AI-aided Deep Reinforcement Learning for Optical Networks Optimization [41]	AI-generated policy design paradigm for optical networks	Representative 14-node NSFNET topology. Analysed 2 NP-hard problems (RWA, RMSA).	Achieved the highest reward and lowest blocking rate on RWA and RMSA problems across benchmarks.	Integrating advanced generative models into reinforcement learning frameworks offers significant promise for enhancing performance in network optimisation tasks.

Table 6.1: Comparative summary of recent papers on the topic of GenAI for performance management

Deep Generative Model and Its Applications in Efficient Wireless Network Management offers a high-level tutorial and case study, showing how diffusion models can simulate mobile user behaviour and design contracts that optimise wireless service delivery under QoS constraints. This paper sets the foundation for using generative models in performance-aligned incentive design. The authors recommend using diffusion models, GANs, and VAEs to model user demand, spatial-temporal traffic patterns, and mobility trends in environments where labelled data is limited or expensive to collect. Their proposed approach involves training these models on historical network data to learn user behaviour and generate future scenarios that reflect diverse QoS demands. These synthetic samples can then guide contract design, resource allocation, and service-level planning in a way that anticipates usage patterns rather than simply reacting to them.

Deep-Q presents an early but influential application of a deep generative model to directly infer QoS metrics from raw traffic traces. It avoids traditional rule-based modelling and demonstrates that generative networks can accurately reconstruct and understand network performance features from passive observations.

In *Generative-AI for AI/ML Model Adaptive Retraining*, the focus is on sustaining ML model performance over time. Here, GenAI is used to simulate performance degradation, detect relevance drift, and trigger retraining when network behaviour shifts, thus enabling more robust predictive pipelines for 5G slicing.

GAN-powered Deep Distributional Reinforcement Learning addresses performance through dynamic resource allocation in network slicing. It shows how GAN-generated synthetic samples can accelerate policy learning, leading to faster convergence and more stable performance even under traffic fluctuation. The idea presented in the paper is that the GAN can learn and then generate realistic synthetic data. The synthetic samples improve the sample efficiency of the learning process and help the DRL agent form more robust policies early in training. In this way, better adaptation to variability in traffic and service demand can be achieved, which is important in 5G slicing, where conditions change rapidly and decisions must be made with incomplete or delayed feedback.

OpticGAI targets optical networks, combining generative learning with deep reinforcement techniques to optimise complex problems like routing and spectrum allocation. It achieves strong empirical results, including reduced blocking rates and better resource efficiency, illustrating GenAI's ability to guide real-time decision-making under performance constraints.

Two common threads emerge across these models: predictive performance tuning and network simulation. Generative models are used to anticipate future behaviour (e.g., retraining triggers, user contract outcomes) and simulate hard-to-observe conditions such as rare loads.

From a practical standpoint, the advantages include support for proactive scaling, more informed load balancing, and enhanced decision-making under uncertainty. However, trade-offs remain. The models require high-quality, representative training data, particularly for traffic synthesis and dynamic policy learning. Additionally, their computational overhead must be carefully managed to ensure their use in production environments.

6.3 Strengths and Open Challenges

Generative AI shows clear strengths in addressing performance management tasks that involve prediction, adaptation, and optimisation under dynamic network conditions. In particular, generative models are used to anticipate traffic loads, simulate user behaviour, infer quality metrics, and support intelligent decision-making frameworks.

One of the most consistent advantages is support for proactive performance tuning. Time-series generative models and diffusion-based techniques forecast conditions such as traffic spikes, enabling systems to pre-empt congestion and adjust resource allocation before degradation occurs. This shifts performance management from reactive intervention to anticipatory control.

Another strength is the ability to support complex decision-making under uncertainty. In reinforcement learning settings, GAN-generated synthetic samples help accelerate training and improve policy stability for problems such as network slicing and spectrum assignment. This leads to more efficient resource use and better adherence to performance guarantees.

However, several open challenges remain. One is the need for high-quality training data, both in terms of representativeness and structure. Many generative models rely on detailed traffic traces or labelled performance metrics, which are not always readily available or may require extensive preprocessing.

Another issue that needs to be taken into account is computational cost. While generative methods have demonstrated strong performance in simulation and training, their deployment in real-time network environments must contend with latency, scalability, and energy efficiency constraints.

Finally, the lack of standardised benchmarks and evaluation scenarios makes it difficult to compare results and generalise the findings to diverse network settings. As the field matures, more systematic validation and cross-environment testing will be needed to ensure the reliability and transferability of generative approaches to performance management.

7 Security Management

Within the FCAPS framework, security management involves safeguarding every network layer against threats and unauthorised access. Its scope spans users and identities, data in motion and at rest, network and compute infrastructure, and the operational telemetry that drives incident response. However, it also ensures confidentiality, integrity, availability and authenticity across enterprise, cloud, edge, and industrial IoT environments.

Generative AI now expands how this mandate can be met. By learning the fine-grained patterns of normal and malicious behaviour, generative models can support a variety of cybersecurity applications, such as:

- Creating realistic attack traffic for testing and evaluation purposes.
- Crafting detection rules directly from high-level threat descriptions.
- Powering anomaly-detection engines that surface novel intrusions with fewer false alarms.

These capabilities move security management from reactive, rule-driven processes toward proactive, adaptive defence, while also introducing new challenges around model robustness, privacy and governance.

7.1 Opportunities for Generative AI

Generative AI instantly transforms security management by allowing defenders to imagine attacks before they happen and translate high-level threat knowledge into machine-enforceable defences. Because models such as GANs, VAEs, diffusion transformers and LLMs can learn the fine statistical structure of network traffic, log events and malware artefacts, they are uniquely suited to two long-standing pain points: (i) the chronic shortage of labelled examples for rare or never-before-seen intrusions, and (ii) the slow, error-prone process of turning prose-level intelligence into concrete detection logic.

The first opportunity lies in **synthetic attack generation**. A generative model trained on even a handful of real intrusions can emit complete packet captures, NetFlow records, or multivariate sensor traces that convincingly reproduce timing, size, protocol mix and payload quirks while varying non-critical fields so the output is not a carbon copy of the source data [42] [43]. Security teams replay these synthetic traces through sandboxes, continuous-integration pipelines or replica networks to stress-test intrusion-detection and prevention systems under conditions that rarely occur in production logs: low-and-slow reconnaissance, polymorphic data exfiltration, multi-stage lateral movement, bursty DDoS floods, or coordinated IoT botnet surges. Because the generator labels every packet or log line as they are produced, it delivers perfectly annotated datasets that balance minority classes without exposing live infrastructure to risk. In operational environments, this practice has already cut weeks out of model-retraining cycles and revealed blind spots that signature-only defences had missed, especially in cloud, 5G core, industrial-control and smart-city deployments where the attack surface changes faster than policies can be written by hand.

The second opportunity is **prompt-based rule generation**. Large language models can ingest a brief, plain-English threat description and optional context, such as PCAPs (packet-capture files that record raw network traffic), log excerpts, or asset inventories, and emit ready-to-deploy detection code for multiple enforcement layers. For deep-packet-inspection engines like Suricata and Snort, the model writes rules that define protocols, offsets, byte patterns, flow direction, thresholds and metadata in each engine's native, keyword-rich syntax. Runtime and file-integrity monitoring can generate YAML-based policies whose conditions, exceptions and severities are declared concisely and version-controllable. When the goal is centralised log analytics, the same prompt can yield Sigma rules, an expressive, vendor-neutral language understood by Security Information and Event Management (SIEM) platforms. Crucially, the LLM grounds its output in the organisation's environment: it substitutes internal IP ranges, proxy ports, domain names, and cloud-specific artefacts so that each rule fires only where it should, and it appends a short human-readable rationale for peer review. What once took an analyst an afternoon of syntax look-ups and test-deploy cycles can now be accomplished in minutes, with the model iteratively refining the rule as fresh telemetry arrives.

The LLM may even suggest playbook updates for the incident-response team. The result is an agile security posture in which detectors, rules and response procedures evolve almost as quickly as the threats themselves, constrained by governance layers that validate model output, check for overfitting or hallucination, and enforce privacy boundaries. In this way, generative AI shifts security management from a reactive, manual discipline to a proactive, adaptive system that anticipates attacks, prepares tailored countermeasures and continually hardens itself with every cycle.

7.2 Review of GenAI Techniques and Models

GenAI research in security management centres on two complementary goals. The first is **data augmentation**: models such as Generative Adversarial Networks (GANs)—pairs of neural nets where a generator fabricates traffic and a discriminator tries to spot the fakes—produce realistic, pre-labelled attack flows that balance otherwise sparse classes. The second is **unsupervised anomaly detection**: models like Variational Autoencoders (VAEs), which compress data into a probabilistic latent code (i.e., a compact representation where each feature is modelled as a probability distribution rather than a fixed value) and reconstruct it, or Long Short-Term Memory (LSTM) recurrent networks that track temporal patterns, learn a concise description of "normal" behaviour and flag any deviation. Recently, transformer-based language models—attention-driven networks initially designed for natural-language tasks—have been adapted to generate intrusion-detection rules or summarise alerts. In practice, these components are often chained: a GAN fabricates botnet traffic to enrich the training set, while its discriminator, or a companion autoencoder, operates online as the anomaly sensor. Increasingly, the entire pipeline runs inside privacy-preserving or edge-aware frameworks so raw telemetry never leaves the device, aligning with data-sovereignty mandates and the inherently distributed nature of modern networks.

Reference	Approach	Experiment Setup	Success Level	Key Findings
Generative AI in Network Security and Intrusion Detection [44]	Comprehensive survey of GANs, VAEs and diffusion models for anomaly detection in DoS/DDoS attacks, behaviour control, and deep-packet inspection.	No new experiments – synthesises ~120 prior GAN/VAE/Diffusion NIDS studies drawn from real enterprise, IoT and ICS deployments.	Qualitative synthesis: the study maps research progress and open challenges.	Generative models excel at zero-day intrusion discovery and data-scarce environments. It warns of adversarial attacks on the models themselves and calls for robust, explainable GenAI defences.
An Enhanced AI-Based Network Intrusion Detection System Using GANs [42]	WGAN produces realistic DoS/DDoS flows; the auto-encoder stack detects anomalies on the augmented data.	Benchmarked on NSL-KDD (41 features), UNSW-NB15, a proprietary IoT trace, and a small real-traffic capture, 70/30 train/test splits with five-fold cross-validation.	93.2% accuracy (NSL-KDD) and 87% (UNSW); minority-class recall ↑ 16 – 24% over baseline CNN/RF models.	GAN-generated traffic balances class distributions, sharply raising recall on rare IoT attacks while keeping the false-positive rate below 2%.

Reference	Approach	Experiment Setup	Success Level	Key Findings
The Identification of Network Intrusions with Generative AI Approach for Cybersecurity [45]	Compares GAN, stacked auto-encoder and SVM on identical feature pipelines in DoS/DDoS, R2L and U2R attacks.	NSL-KDD dataset; 80/20 split; identical preprocessing (missing-value handling, one-hot, min-max scaling).	GAN reached 91.12% accuracy, beating VAE (88.5%) and SVM (84%).	Generative modelling captures high-dimensional attack structure better than purely discriminative baselines, showing higher TPR on U2R and R2L classes.
On the Usage of Generative Models for Network Anomaly Detection in Multivariate Time-Series [46]	Net-GAN (LSTM-GAN) & Net-VAE for unsupervised multivariate traffic series.	IoT sensor network, enterprise NetFlow logs; models trained on normal traffic, then evaluated on injected botnet, port-scan, DDoS, and ICS attack traces.	Botnet 93%, port-scan 89%, DDoS 78% at <1% FPR; 70% of ICS attacks caught with zero FPR.	Temporal generative modelling slashes false positives while retaining high detection on subtle multi-metric anomalies.
MAD-GAN: Multivariate Anomaly Detection for Time-Series Data with GANs [47]	LSTM-GAN with dual reconstruction + discrimination "DR-score" for unsupervised CPS anomaly detection.	SWaT (51 sensors) and WADI (123 sensors) water-treatment testbeds; trained only on normal weeks, evaluated on 34 attack scenarios.	Detects all staged attack types, outperforming PCA and variational-RNN thresholds by 12-20% in F1.	Joint modelling of all sensor channels lets MAD-GAN spot cross-variable drift invisible to univariate or threshold methods, giving robust ICS protection.
Intrusion Detection Using Distributed Multi-Level Discriminator GAN for IoT Systems [48]	Federated GAN: each IoV/IoT node trains a local LSTM-GAN; hierarchical discriminators fuse decisions without sharing raw data.	Simulated 50-node IoV; KDD99 & SWaT used as traffic templates; edge-level training with periodic model aggregation.	Accuracy 98.9%, recall 98.98%, F1 0.77—outperforms centralised GAN and CNN baselines; latency < 210ms per batch on Raspberry Pi-4 class devices.	A multi-discriminator federation yields a privacy-preserving, scalable IDS that maintains high detection even with highly imbalanced DoS/spoofing traffic.

Table 7.1: Comparative summary of recent papers on GenAI for security management

Recent literature exhibits three clear trends. First, almost every study tackles class imbalance by synthesising rare intrusions—slow data exfiltration, infiltration beacons, stealthy port scans—so that supervised detectors no longer bias toward dominant DoS traffic. Second, models are becoming temporal and multivariate: RNN-based GANs and VAEs capture cross-metric correlations in IoT plants and multi-sensor fleets, slashing false-positive rates to below one per cent. Third, researchers are moving the computation toward the edge; distributed discriminators and federated training let cameras, vehicles or programmable logic controllers (PLCs) defend themselves without exporting sensitive packets.

Across the corpus, several strengths recur. Generative augmentation reliably lifts recall for minority attacks by 5–12 percentage points [42], and unsupervised variants detect zero-days that signature sets miss, all while maintaining operationally acceptable alarm rates—the same generators double as Red-Team engines[2], crafting evasive payloads that harden blue-team models in self-play. Finally, prompt-controlled language models are

[2] A Red Team is a team of cybersecurity experts that imitates network assaults on a company to find weaknesses. The Blue Team, on the other hand, is in charge of preventing assaults and maintaining the security posture of the company.

emerging that turn natural-language threat briefs into Sigma [49] or Suricata [50] rules in minutes, collapsing the lag between intelligence and enforcement.

Yet limitations persist. Most evaluations remain offline or on replay testbeds; throughput and latency under live traffic are rarely profiled. GAN training is compute-heavy and can collapse without careful tuning, while VAEs may underfit highly volatile cloud workloads. None of the surveyed detectors is yet adversarially robust—attackers can, in principle, train their own GANs to generate traffic that blinds the model. Governance tooling is also thin: only a handful of papers validate that synthetic data cannot be reverse-engineered to leak user secrets, and policy-aware rule generation still depends on expert review.

The work shows that generative AI can raise detection accuracy, broaden coverage to unseen threats and automate rule synthesis, provided organisations invest in reliable training pipelines, runtime efficiency and strong validation layers. Those guardrails, privacy-preserving deployment patterns, and adversarial hardening remain the critical research frontier as generative security moves from prototype to production.

7.3 Strengths and Open Challenges

GenAI approaches bring several clear advantages to security management. First, they improve detection accuracy in situations where labelled data is scarce. By producing high-fidelity synthetic traffic, a GAN or diffusion model can balance datasets containing only a handful of examples of slow data exfiltration, stealth scans or niche IoT malware, allowing supervised detectors to learn those patterns instead of ignoring them. Second, unsupervised variants—such as VAEs or LSTM-driven GAN discriminators—create a precise statistical portrait of normal network and sensor behaviour. When a slight deviation occurs, they raise an alert with far fewer false positives than signature-based systems, a critical benefit in operational technology and industrial IoT environments where excessive alarms can halt production. Third, large language models accelerate the analyst workflow by translating plain-English threat intelligence into machine-readable rules for engines like Suricata, Snort or Sigma, shrinking tasks that once took hours to minutes. Finally, these generative and discriminative components can be joined in a self-improving loop: synthetic attacks probe the stack, missed detections feed fresh training data, and updated rules roll out automatically, creating an adaptive defence that evolves almost as fast as the threats themselves.

However, generative AI also introduces open challenges that must be addressed before these systems can be considered production-ready at scale. Most published evaluations are still confined to replay testbeds or laboratory traffic; little is known about generative detectors' throughput, latency and failure modes running on busy enterprise backbones or low-power edge devices. Training GANs is computationally expensive and notoriously unstable, while VAEs can underfit highly volatile cloud workloads unless hyperparameters are tuned with care. Privacy and governance pose further hurdles: synthetic traffic must avoid leaking sensitive patterns, and automatically generated rules must be reviewed for hallucination, policy conflicts or overfitting to a narrow slice of data. Adversarial robustness remains unexplored, mainly because an attacker with their own GAN might craft network flows that fool the detector trained to stop it. Finally, integrating these models into existing security operations pipelines requires precise version control, rollback and auditing mechanisms so that automated updates never compromise compliance requirements.

Addressing these weaknesses—through rigorous runtime profiling, privacy-preserving training, adversarial testing and strong human-in-the-loop validation—will determine whether generative AI matures from a promising research direction into a dependable front-line defence.

8 AI Tools for Network Management

Tools based on GenAI are quickly emerging and becoming more relevant in the area of network management. This section aims to give an overview of the key tools available as of Summer 2025 and outline trends in the use of GenAI for network management. As with the research reviews in previous sections, this section maps the tools to the well-known ISO FCAPS management categories. This mapping is not perfect, as some tools cover more than one category, and the boundaries between them are not always clear, but it gives a good overview of where a tool is positioned from the perspective of network management.

AI-based network management tools have many novel features, such as script generation through the use of natural language, the reduction of manual intervention, and performance and event analysis. This means that complex tasks can be solved quickly, without intensive prior knowledge. The exponential growth of traffic, variability in services, complex topologies, and the adoption of distributed computing architectures are pushing the limits of conventional tools. In this complex landscape, GenAI emerges as a transformative force for network management solutions through natural language interaction, enabling network engineers to query network performance information and status data, which can be interpreted by tools to provide understandable answers and insights. Additionally, these tools can proactively suggest network adjustments or optimal configurations to prevent performance degradation before it occurs, according to baseline configurations.

This section provides an overview of both open-source and commercial tools within the fields of AIOps (Artificial Intelligence for IT Operations) and network performance management. The primary differences between the two categories, open-source and commercial, lie in their development models, costs, ease of implementation, support, and required level of expertise. Commercial GenAI tools are primarily integrated as advanced features within existing vendor AIOps and network performance management platforms. In general, commercial GenAI tools have been developed to optimise processes and workflows in a network to save resources, allowing employees to spend less time on network management tasks and on optimising resources and devices to boost their efficiency.

In most cases, commercial vendors advertise statistical analyses for all the improvements and savings (in both money and time) that can be achieved with their tools compared to not using them. They usually offer ease of deployment, dedicated customer support, and maintenance. Open-source tools, on the other hand, often require some degree of configuration and expertise to implement and utilise effectively. While open-source tools offer many customisation possibilities, they often lack the support offered with commercial solutions. In general, there are not a huge number of open-source network management tools – those that exist are typically provided free of charge by developers and require more effort to fully implement in a network environment compared to ready-made commercial software.

Existing GenAI tools in network management are usually focused on a combination of the Fault, Configuration, and Performance management domains. This is driven by the operational interdependencies between these domains and the aim of adopting an autonomous network paradigm. For instance, a GenAI tool can predict an upcoming performance degradation by analysing real-time metrics and diagnosing it as a possible fault. Subsequently, the tool can automatically generate and implement an optimal configuration change to prevent the fault or mitigate its impact. As most solutions are designed to support these critical roles, Section 8.1 analyses tools that follow a joint Fault, Configuration, and Performance management approach. However, it should be noted that they can also be focused on security, and directly or indirectly on accounting management.

Another clear FCAPS management category that is supported by GenAI tools is Security management, in the context of the ever-growing threat of complex cyber attacks pushing companies to invest more in their cybersecurity. As a result, more and more commercial GenAI tools for security management are emerging. ML-based tools and methods have also been in use for some time for anomaly and intrusion detection. Although security is not the primary focus of GN5-2 Work Package 6, due to the importance of the risks involved in network security and the number of existing tools, security management tools are described separately in Section 8.2.

8.1 Fault, Configuration, and Performance Management

The following table presents a selection of tools categorised by their core contributions to the FCAPS network management domains of configuration, performance, and fault management. The tools are presented based on the generative AI features they utilise, such as natural-language generation, code synthesis or threat detection. References are included to support further exploration and validation. The licensing model for a tool is labelled in the C/O column, with 'C' indicating ‘Commercial’ for proprietary software and 'O' for ‘Open Source’, publicly available and freely distributable code. This overview highlights how generative AI is being applied across diverse network operational areas to enhance automation, decision-making, and efficiency, supporting human operators to enable proactive and intelligent orchestration.

Most tools offer similar solutions, such as troubleshooting assistants, anomaly detection and script generation. These troubleshooting assistants employ generative AI to analyse network logs, telemetry data and historical records, identifying the root cause of issues, suggesting potential solutions, and guiding humans through step-by-step resolution processes, thereby significantly reducing downtime. Anomaly detection, another common feature, leverages continuous monitoring of network behaviour to learn standard patterns and identify unusual activities or performance deviations that might indicate upcoming issues, often before they impact users. The tools often involve script generation, creating configuration scripts that automate manual processes to facilitate faster deployment and consistent network configurations.

GenAI tools typically do not have dedicated accounting features, as this strength lies in analysing amounts of network data for patterns, anomalies, and predictions related to fault, configuration, and performance management. However, by providing insights into resource utilisation, identifying inefficiencies, and forecasting capacity needs, GenAI tools indirectly contribute to accounting by enabling resource optimisation in network operations.

In addition to commercial tools, a specific set of solutions has arisen that leverages open-source, code-generating LLMs, indicating a significant maturation beyond general-purpose text generation. These models are mainly tailored to produce accurate network configurations and automation scripts. This directly addresses the critical need for reliable automated network programming.

Tool	FCAPS	Generative Features	Description	C / O
Cisco AI Network Analytics / Generative AI Policy Assistant [51]	F/C/P/S	• AI-driven root cause analysis and fault prediction • Automated network configuration • Performance anomaly detection • Natural-language query interface • Security threat detection	Cisco AI Network Analytics is an application embedded in Cisco DNA Center that leverages its capabilities to generate AI-driven insights for devices, applications and networks, guaranteeing data protection and privacy. Automates repetitive tasks, identifies problems at an early stage and suggests remediation in the areas of performance, fault, security, and configuration.	C
Juniper Mist AI / Marvis AI Assistant [52]	F/C/P/S	• AI-driven anomaly detection • Natural-language querying • Automated network configuration • Proactive fault detection	Leverages LLMs and NLP to power the Marvis AI Assistant. This feature enables users to interact with the network with conversational language for troubleshooting, configuration assistance, and network insights. The tool generates reports, summarises network status, and suggests configurations based on user inputs.	C
Fortinet / FortiManager / FortiAnalyzer / FortiMonitor / FortiAI [53]	F/C/P/S	• Script generation • SD-WAN • Troubleshooting and LLM querying • Policy creation • IoT device analysis • Anomaly detection	Uses input data such as device configurations, security policies, logs, network topology, and user prompts to automate and optimise firewall and network management. It also integrates FortAI to recommend policy adjustments and eliminate repetitive efforts. Aids in generating scripts, resolving issues, and optimising security operations via natural-language interactions.	C
IBM's Watsonx [54]	F/C/P/S	• Advanced LLM-based natural language processing • Conversational AI • AI-assisted root cause analysis • Automated network configuration suggestions • Anomaly detection • Security threat detection	A portfolio of AI products that accelerates the impact of GenAI in core workflows to drive productivity, offering flexibility, trusted outputs, and access to structured / unstructured data. Delivers AI-powered insights through techniques such as retrieval augmented generation and fine-tuning.	C

Tool	FCAPS	Generative Features	Description	C / O
NetBrain's Self-Service Chatbot [55]	F/C/P/S	• Natural-language network troubleshooting • LLM to streamline user interactions • Natural-language processing	Uses user-input natural language, combined with a no-code library of pre-built network automation units to manage infrastructure and diagnose and resolve issues. Retrieves real-time data from network devices through APIs, uses dynamic topology maps for reference, and verifies against established configuration baselines. The chatbot provides automated, context-sensitive troubleshooting and remediation via a conversational interface.	C
Atera's AI-Powered IT Management Platform (AI CoPilot Assistant) [56]	F/C/P/S	• Script generation • Troubleshooting assistant • AI ticket triage • Anomaly detection • Remote access security	Uses input data such as device health metrics, ticket histories, remote session logs, knowledge base articles, endpoint security info and natural-language prompts. It utilises this data to produce scripts, identify problems, and automate ticket processing.	C
ServiceNow IT Infrastructure Management [57]	F/C/P/S	• Natural-language interaction • Article generation • Automated incident summarisation • Policy and workflow generation	Provides real-time monitoring, predictive analytics, and intelligent automation to optimise infrastructure performance, detect anomalies, and proactively address potential issues. The tool collects device input data from multiple sources, such as network discovery protocols (SNMP, WMI, SSH), its Configuration Management Database, integrations with third-party monitoring tools, and telemetry data from network devices.	C
OpenAI Codex* [58]	F/C/P/S	• Natural-language processing • Reads and edits files • Runs commands	Built on a fine-tuned version of GPT-3 to understand and generate code in various programming languages. It acts as an autonomous code assistant, as the code can be debugged, refactored, and tested within sandboxed environments.	C
Broadcom DX Operational Observability [59]	F/P	• Generative AI summarisation for services • Natural-language filtering	Simplifies the understanding of complex incidents, makes data filtering intuitive for non-technical users, and accelerates the triage process.	C
XenonStack GenAI for NOC [60]	F/P/C	• Anomaly detection	Gathers data from applications, infrastructure, network logs, metrics and events to train ML models and detect anomalies.	C

Tool	FCAPS	Generative Features	Description	C / O
NetGPT [61]	F/C/P	• Natural-language troubleshooting assistant • Device interaction chatbot • LLM interface	NetGPT employs a collaborative cloud-edge methodology, optimising the deployment of LLMs based on their computational capacities, allowing efficient processing of network traffic data.	O
LangChain* [62]	F/C/P	• Script generation • Natural-language device configuration	Designed to build applications powered by LLMs, including AI assistants, internal copilots, and automation tools across network management domains. Tested in multiple studies enabling interaction with 3GPP documentation and supporting generation and structuring of dataset metadata [63], along with a LangChain-based network management assistant that adapts to network baselines [64], demonstrating utility not only in operational monitoring but also documentation, compliance and decision-support.	O
PydanticAI* [65]	F/C/P	• Agentic orchestration • Streaming response generation	Supports different models (e.g., Gemini, OpenAI, Ollama). Leverages Pydantic models to define and enforce the output schema of an LLM with a model-agnostic architecture.	O
Google's Agent Development Kit (ADK) [66]	F/C/P	• Multi-agent orchestration • Tool-based reasoning and execution • Agent-to-agent communication protocols	A Python toolkit for building and evaluating multi-agent AI systems through their step-by-step execution trajectories. Enables agents to collaborate on complex tasks.	O
MS AutoGen* [67]	F/C/P	• Multi-agent conversation and collaboration • Code execution • Autonomous problem-solving	Allows the creation of multiple AI agents to communicate and collaborate to solve complex problems. Supports enhanced LLM inference APIs.	O

* Tool not designed as a direct network management tool. However, it is a comprehensive framework for developing AI-powered applications that can be used for network management.

Table 8.1: Classification of fault, configuration, and performance management tools based on GenAI

Cisco AI Network Analytics, Generative AI Policy Assistant

Cisco AI Network Analytics [68] is an application embedded into Cisco DNA Center. The tool focuses on intelligent issue detection and analysis to enhance the network performance management functionalities of Cisco DNA Center, identifying network events and sending them to the Cisco cloud. AI-driven baselining is employed to analyse the network behaviour of a specific network environment, learning the baseline network behaviour and using it as a reference to identify performance issues. Unusual network patterns and their causes (e.g., connection issues, application experience issues) are identified by an AI-driven anomaly detection feature. Furthermore, the tool offers comparative analytics by comparing a specific KPI from the analysed network with another network.

Another powerful tool relying on generative AI is Cisco AI Assistant [69], which utilises conversational AI for IT and security operations. The assistant leverages its capabilities to generate AI-driven insights for devices, applications and networks, guaranteeing data protection and privacy. The user can ask questions and request information such as "Give a person access to a specific app", "Show me WAN port usage", "How can I troubleshoot the alert "DHCP no leases"?", or "What are the group policies currently configured?".

By automating repetitive tasks, identifying problems at an early stage and suggesting measures in the areas of performance, security, and configuration, Cisco AI Assistant improves efficiency, reliability and protection while enhancing network intelligence.

The tool applies to the following FCAPS management domains:

- **Fault management:**

  It helps to simplify root cause analysis using a conversational interface for troubleshooting.

- **Configuration management:**

  The assistant can automate a variety of setup tasks to reduce routine network management processes. It provides proactive suggestions to enhance network reliability and stability, and helps users understand and manage group policies, client policies, and network-wide policy configurations.

- **Performance management:**

  It provides real-time visibility and understanding of the network and suggestions to enhance network reliability.

- **Security management:**

  The assistant aids in troubleshooting security appliances, firewall rules, VPN configurations, and network security events.

Juniper Mist AI

Juniper Mist AI [70] is a tool designed for AI-driven operations, using AI at the core. The platform enables advanced capabilities, real-time optimisation, and autonomous issue resolution. Its features include data collection from telemetry and user state and decision-making.

The tool utilises conversational and generative AI capabilities to provide proactive insights, explain anomalies, and suggest potential fixes in natural language, enhancing operators' troubleshooting capabilities. Network performance can be monitored to identify bottlenecks and address usage issues. Mist AI uses a cloud-native, microservices-based architecture that is scalable, open, and API-driven.

The GenAI-based Marvis AI Assistant [71] is integrated with the Juniper Mist AI platform, providing a chat interface with real-time responses and acting as a co-pilot to respond to issues. Marvis AI Assistant draws on historical information from Juniper's public-facing knowledge base and employs NLP to optimise experience in wired and wireless access, SD-WAN, and WAN domains. Agentic AI capabilities are provided by leveraging multiple agents and user feedback. The assistant can answer direct questions about network state, recommend proactive actions, and provide support for configuration management. This serves to reduce mean time to resolution (MTTR), reduce human effort by correlating data to derive conclusions, and ensure a seamless and smooth user experience. The tool offers a dashboard view of high-impact network issues at an organisational level.

The tool applies to the following FCAPS management domains [72]:

- **Fault management:**

  The tool provides automated root cause analysis, intelligent alert correlation, a conversational interface for troubleshooting, and automated remediation script generation/execution.
- **Configuration management:**

  Juniper Mist AI supports self-configuring networks for faster deployment, and Agentic AI can continuously monitor and adjust network configurations to optimise performance, security, and resource utilisation in real time.
- **Performance management:**

  The tool enables network administrators to define SLAs and monitor them in real time.
- **Security management:**

  It offers integrated security and control, with a unified Mist dashboard delivering Zero Trust at scale through consolidated management and consistent policy enforcement. Its AI-native security unifies AIOps, networking, and security to provide complete environment visibility and insights into the network implications of security actions. Agentic AI can also proactively identify and neutralise cyber threats.

Fortinet: FortiManager / FortiAnalyzer / FortiMonitor / FortiAI

Fortinet is a cybersecurity company specialising in the protection of corporate networks, data centres, cloud environments and end devices. The company’s software portfolio has adapted in light of the ever-increasing threat of complex cyberattacks, as well as the constant development of GenAI. Numerous GenAI software extensions, including script generation, IoT device analysis and SD-WAN troubleshooting, have created a compact system that can be used in all areas of the FCAPS framework. These features take as input the device configurations, security policies, protocols, network topology, and user prompts to automate and optimise network management. However, most of these features are limited to Fortinet products, especially in terms of configuration, as Fortinet software and hardware are not designed to manage third-party devices nor support other vendors' configurations, protocols, and security models.

The most important Fortinet software extensions apply to the FCAPS framework as follows:

- **Fault management:**

  FortiManager and **FortiAnalyzer** offer capabilities to detect, isolate, and report faults across the network. These systems monitor hardware status, network traffic, and system logs in real time, generating alerts for anomalies, failures, or potential threats. Faults can be automatically logged, prioritised, and escalated based on severity, with an LLM on hand for troubleshooting to minimise downtime.
- **Configuration management:**

  FortiManager offers configuration management simplification, offering centralised control for managing configurations across various Fortinet devices. From one interface, administrators can create, alter, deploy, and back up configurations via an LLM. This guarantees consistent policy application, and minimises human errors and the need for expert knowledge.

- **Performance management:**

  Extensions like **FortiManager**, **FortiAnalyzer** and **FortiMonitor** support performance management through real-time monitoring, traffic analytics, and reporting features. These systems can track CPU, memory, and bandwidth usage across Fortinet devices, and monitor network latency and uptime. The system then uses this data to optimise network routing with SD-WAN performance metrics and identify traffic bottlenecks or underperforming links.

- **Security management:**

  FortiManager, **FortiAnalyzer** and **FortAI** provide continuous monitoring, threat assessment, policy implementation, and incident management to protect infrastructure against cyberattacks like malware, ransomware, phishing, and zero-day exploits. Administrators can define and apply security policies consistently across all Fortinet devices from a single management platform with natural language via an LLM. The system constantly observes network traffic and device activities to detect threats immediately, automatically initiating responses like blocking harmful traffic or isolating compromised systems.

8.2 Generative AI Tools for Security Management

The expansion of GenAI is not only affecting application development but also allowing the production of more sophisticated hacker and attack tools. This leads to AI-driven attacks and an expanded attack surface from cloud to IoT devices. To counter this, AI-based security tools are being developed. This means that the market is developing towards countering AI with AI. These tools are designed:

- To protect generative AI systems from attacks and misuse such as prompt injections (malicious inputs aimed to manipulate AI responses) [73].
- To analyse massive volumes of data rapidly, automatically generating threat reports and summaries.
- To simulate attacks to test the defences of networks (Red-Teaming).

It is also possible to enhance existing security tools with generative AI – for example, a network traffic analyser with a dedicated log output could be used to configure an LLM through the Model Context Protocol (MCP)[3]. An MCP server is a central system that manages and distributes context information between LLMs, AI agents or tools. It enables structured, real-time communication and context sharing, often used in multi-agent systems, AI orchestration, or tool-using environments.

This section describes some of the available AI-based tools used not just to enhance network security, but also to strengthen AI models and applications through the AI output, reasoning, and action validation. There are also overlaps with other FCAPS management domains, especially with Performance management, because in most cases, security problems can be analysed using network data or data traffic in general, such as in the case of a DDoS attack. Performance management in GenAI-based security tools focuses on maintaining fast and scalable analysis of large security data streams. This ensures that models deliver accurate and reliable threat detection. Therefore, commercial GenAI security tools depend somewhat on the Performance management domain.

But there is a significant difference with these tools: firstly, there are tools with limited LLM interaction where users can only ask monitoring questions – “What happened in my network and what are your suggestions?” Other tools offer the possibility to configure the network directly with LLM input to a limited extent, i.e., “I have a security problem here, do something about it!” These tools are marked in the FCAPS column in the following table with “S” or “S/C” respectively.

It is impossible to cover all generative AI tools for security management in one table, as this field is one of very active current development. Thus, Table 8.2 lists notable examples of commercial and open-source security management tools that underline the market’s direction towards LLMs.

[3] The Model Context Protocol (MCP) is a communication protocol designed to manage and share contextual information between AI models, tools, or systems in a structured, standardised way.

Tool	FCAPS	Generative Features	Description	C / O
Microsoft Security Pilot [74]	S/C/P	• Natural-language querying and configuration via LLM • Real-time malware analysis	Uses data from organisation's security tools, Microsoft's global threat intelligence, and context-specific information via secure plugins to analyse and respond to threats.	C
Charlotte AI [75]	S/C/P	• Natural-language querying and configuration via LLM • Multi-agent architecture • Threat detection and response	Uses input data from three main sources: Falcon platform telemetry (e.g., endpoint and cloud activity), CrowdStrike threat intelligence (tracking adversaries and attacks), and expert-validated content from services like Falcon Complete and OverWatch. These inputs provide real-time, high-fidelity context for accurate threat detection and response.	C
Darktrace ActiveAI Security Platform [76]	S/P	• Natural-language querying via LLM • Threat detection and response • Self-learning AI	Utilises input data from various sources within the enterprise, such as network traffic, endpoints, cloud services, email systems, identity management systems, and third-party integrations. Builds behavioural baselines and detects anomalies using self-learning AI on this telemetry.	C
Zscaler Zero Trust Exchange [77]	S/P	• Natural-language querying via LLM • Prompt classification and inspection • Threat detection and prevention	Uses a wide range of telemetry, including network, application, device, and SaaS traffic as input data. Incremental training of AI models for policy recommendations, breach prediction, and secure generative AI controls is performed using these inputs in conjunction with external security and business systems.	C
SentinelOne Purple AI [78]	S/C/P	• Natural-language querying and configuration via LLM; multilingual support • Threat hunting • Autonomous investigations	Uses real-time telemetry including endpoint, cloud, identity, and network security events across both native and third-party log sources. It also includes threat intelligence, automated neural-network-based detections, and enriched metadata, which power its generative AI capabilities for natural-language threat hunting and investigation.	C
Tenable ExposureAI [79]	S/P	• Natural-language querying via LLM • Attack path summarisation • Specific mitigation guidance	Uses input data from the Tenable Exposure Graph, including over 1 trillion exposures, configuration findings, assets, and vulnerability data across IT, cloud, and operational technology environments. This data powers its generative AI to support risk analysis, search, remediation workflows and attack path summarisation.	C

Tool	FCAPS	Generative Features	Description	C / O
Coralogix [80]	S/P	• Natural-language querying via LLM • Real-time observability • AI safety focus (including prompt injections & hallucinations)	Utilises input data derived from logs, metrics, traces, and security telemetry from various applications and cloud infrastructure, along with vectorised observability records, model evaluation metadata and real-time security alerts to facilitate semantic search, anomaly detection and AI governance.	C
Lakera Guard [81]	S/P	• Natural-language explanations (no LLM) • Prompt injection protection • Real-time threat detection • Red Team simulations	Ingests all LLM interactions, including user inputs, system prompts, and model outputs, passing them through detectors for prompt injection attacks, jailbreaks, content violations and malicious links. It uses continuous Red Teaming, where data is used to train and update its detection models, enabling real-time protection against evolving GenAI threats.	C
Calypso AI [82]	S/C	• Natural-language querying and configuration (scanners) via LLM • Real-time adaption of scanners • Red Team simulations	Monitors and scans user prompts and AI responses in real time to identify risks such as prompt injections, Personal Identifiable Information (PII) leaks, toxic content and sensitive business data. It employs both built-in and customisable scanners via LLM. It supports Red Teams within a controlled environment to test LLM vulnerabilities. These data help refine custom scanners and strengthening model defences against real-world threats.	C
Lasso Security [83]	S/C	• Cybersecurity solution for LLM • Securing and monitoring AI interactions • Red Team simulations	Examines the prompts and outputs of LLM interactions to identify sensitive data, implement security measures, and avert risks such as PII exposure or IP leaks. It also supports automated red teams by simulating adversarial attacks to identify vulnerabilities in AI workflows.	C
LlamaFirewall [84]	S	• Prompt injection protection • Monitoring agents in real time • Insecure code prevention	Examines all data traversing GenAI pipelines, including user prompts, external content, agent reasoning and produced outputs, to prevent security threats such as prompt injections, goal hijacking, and insecure code. It is completely open source and does not require purchasing API access for any LLM to use the core features. However, there are some community extensions that rely on OpenAI's moderation API, which does require an OpenAI API Key, but this is completely optional.	O
Adversarial Robustness Toolbox (ART) [85]	S	• Ensures robustness against adversarial inputs for generative models (such as evasion, poisoning, extraction and inference attacks)	Supports a wide range of input data types, such as images, audio, video, and text. Its purpose is to assess and improve the robustness of models by creating adversarial examples and evaluating defences within different machine learning frameworks. It is completely open source and free to use.	O

Tool	FCAPS	Generative Features	Description	C / O
Garak [86]	S	• Focus on LLM security: prompt injections, jailbreaks, etc. • Red Team simulations	Employs crafted adversarial prompts to evaluate GenAI models for weaknesses such as hallucinations, prompt injection, and data leakage. Examines the model's replies to detect flaws and enhance Red-Team strategies. It is an open source tool, but offers a commercial enterprise version that includes advanced features like managed Red-Team services, enterprise-only probes, compliance reporting and historical benchmarking.	O
Zeek (Enhanced with AI) [87] [88]	S/P	• Network traffic analyser with detailed log output • Can be used as input for an MCP server for LLM interaction	Uses network traffic data as its primary input, analysing it to generate detailed logs about connections, protocols, files, and other artefacts. It can also ingest external structured data (e.g., logs or tables) to enhance detection and correlation capabilities. This detailed log output can be used to configure an LLM through an MCP server. Zeek is open source, but to use it together with an LLM or to build an MCP server, purchasing API access to an LLM may be required.	O
Cybersecurity AI [89] [90]	S/P	• Red Team simulations • Modular agent design	To simulate attacks, it requires input data like details of the target system, network traffic, vulnerability databases, and tailored exploit scripts. Its outputs consist of reconnaissance reports, logs of exploit attempts, traces of agent activity, and detailed findings from Red Team assessments. It is generally open source, but leveraging external LLMs for scoring, reasoning, or attack generation requires API keys, where purchase may be required.	O

Table 8.2: Classification of security management tools

9 Conclusions

This white paper examined GenAI's rapidly growing role in network management, providing an in-depth analysis of its applications across the FCAPS framework: Fault, Configuration, Accounting, Performance and Security. The review of recently published research showed that GenAI is beginning to reshape operational paradigms across all FCAPS areas.

Configuration and Security management are the most advanced in adoption and capability. LLMs are reliably used in configuration management to convert operator intent into structured, vendor-specific commands. This can significantly reduce the overhead of managing multi-vendor environments and help avoid misconfigurations. In security, generative models such as GANs and diffusion models enable the creation of synthetic attack traces, Red-Team scenarios, and fine-grained intrusion-detection training data. These developments enable more agile, adversary-aware defence mechanisms previously difficult to prototype or test.

Fault management is following closely behind. Research has shown that detection pipelines can be trained more effectively and made more resilient to unseen failure types by using generative models to simulate rare or complex incidents. LLMs are also playing a growing role in this space, assisting with log summarisation, natural-language diagnosis, and co-pilot-style operator interaction. However, challenges remain in validating fault predictions in real time and ensuring they are explainable and actionable.

In performance management, generative models simulate time-series traffic patterns, enable predictive scaling, and optimise resources via generative reinforcement learning. These techniques show strong potential for data centre orchestration, network slicing, and 5G/6G edge optimisation. However, practical deployments are still rare, and the field needs more robust mechanisms for integrating generative outputs into live orchestration pipelines without compromising stability or latency.

Accounting management is currently the least explored FCAPS area regarding generative AI applications. While some underlying techniques, such as automated report generation, summarisation of usage patterns, and policy-aware document drafting, have been demonstrated in related fields like finance and billing automation, dedicated research in the networking context remains limited. Nevertheless, there is a lot of potential, particularly in combining generative models with logs, telemetry, and cost models to streamline chargeback, forecasting, and SLA verification processes.

One consistent theme across all FCAPS areas is that the most effective solutions often rely on hybrid model architectures. For instance, pairing LLMs with GAN-generated training data improves classification robustness, while combining VAEs with LLM prompts enables constrained but human-interpretable configuration synthesis. Similarly, integrating RAG with policy rulebases allows for more factual, verifiable outputs. These architectures enable generative AI to become part of modular pipelines in complex real-world network operations.

However, challenges remain. Generative models can produce plausible but invalid outputs. In a networked system, even a small error in a configuration or a misinterpreted alert can lead to outages, degraded QoS, or security vulnerabilities. The issue of explainability is a problem, especially for GANs and diffusion models, which offer little interpretability out of the box. Furthermore, operational constraints such as latency, integration with other systems, and compliance requirements limit where and how generative AI can be applied in production environments.

Future steps must include developing safeguards and best practices to address these risks. These include human-in-the-loop oversight for critical decisions, post-generation validation layers, training on high-quality and diverse datasets, and using constraint-aware generation techniques. Equally important is the need for benchmarking frameworks and open evaluation datasets that allow results to be compared transparently and fairly across different deployments.

A key consideration is the selection of an appropriate network management solution. The selection requires consideration of stakeholders' specific needs, available resources, and directions. For instance, while open source solutions are often free of charge for developers, they demand higher technical proficiency for successful deployment, customisation, and troubleshooting. Furthermore, the level of support required must be carefully considered. Commercial solutions generally provide dedicated vendor support, whereas open-source tools depend more on community support.

Current market offerings primarily concentrate on real-time performance monitoring, advanced observability, traffic engineering, dynamic resource allocation, and predictive capacity planning. The leading vendors often provide dynamic dashboard generation, unified data insights, proactive resource optimisation, simulations of user connections, dynamic configuration adjustments and performance monitoring in their solutions. These GenAI capabilities enable network configuration to be adjusted dynamically and manually with simple human language via LLMs, optimise traffic flow and forecast future needs, moving the industry toward more autonomous and self-healing operations that meet the escalating demands of modern IT infrastructures and the exponential growth of data.

Generative AI has the potential to shift our perception of automation, transforming it into an adaptive approach based on data-driven systems that can learn and create. Realising this vision requires collaboration between researchers, vendors, operators, and standardisation bodies to ensure that GenAI becomes a trustworthy, sustainable pillar of modern network operations.

Glossary

ACL	Access Control List
AI	Artificial Intelligence
AIOps	Artificial Intelligence for IT Operations
AIGC	Artificial-Intelligence-Generated Content
AP	Access Point
API	Application Programming Interface
ART	Adversarial Robustness Toolbox
BGP	Border Gateway Protocol
BS	Base Station
CDR	Call Detail Record
CL	Convolutional LSTM
CLI	Command Line Interface
CNN	Convolutional Neural Networks
CPS	Cyber-Physical Systems
CPU	Control Processing Unit
DDOS	Distributed Denial of Service
DDQN	Double Deep Q-Network
DHCP	Dynamic Host Control Protocol
DL	Deep Learning
DNA	Digital Network Architecture
DR	Domain Rating
DRL	Deep Reinforcement Learning
ETSI	European Telecommunications Standards Institute
FCAPS	Fault, Configuration, Accounting, Performance, and Security
FPR	False Positive Rate
GAE	Graph Autoencoder
GAN	Generative Adversarial Network
GDT	Generative Digital Twin
GNN	Graph Neural Network
GPT	Generative Pre-Trained Transformer
GSMA	Global System for Mobile Communications Association
IBN	Intent-Based Networking
ICANN	Internet Corporation for Assigned Names and Numbers
ICS	Industrial Control System
IDS	Intrusion Detection System
IP	Internet Protocol
IPR	Intellectual Property Rights
ISO	International Organization for Standardization
IT	Internet Technologies
KDD	Knowledge Discovery and Data Mining
KPI	Key Performance Indicator
LLM	Large Language Model
LSTM	Long Short-Term Memory
LTE	Long-Term Evolution
MAC	Medium Access Control
MCP	Model Context Protocol
ML	Machine Learning

MPLS	Multiprotocol Label Switching
MPPCA	Mixtures of Probabilistic Principal Component Analysers
MS	Mobile Station
MTTR	Mean Time to Resolution
NAT	Network Address Translation
NER	Named-Entity Recognition
NETCONF	Network Configuration Protocol
NIDS	Network Intrusion Detection Systems
NLP	Natural Language Processing
NOC	Network Operations Centre
NP	Non-deterministic Polynomial-time
NR	New Radio
NS	Network Simulator
NSFNET	National Science Foundation Network
OSPF	Open Shortest Path First
OSS	Operations Support Systems
OT	Operational Technology
PCA	Principal Component Analysis
PII	Personal Identifiable Information
PLC	Programmable Logic Controller
QoE	Quality of Experience
QoS	Quality of Service
RAG	Retrieval-Augmented Generation
RAN	Radio Access Network
RF	Random Forest
RIP	Routing Information Protocol
RMSA	Routing, Modulation, and Spectrum Allocation
RNN	Recurrent Neural Network
RWA	Routing and Wavelength Assignment
SD	Software Defined
SIEM	Security Information and Event Management
SLA	Service Level Agreement
SNMP	Simple Network Management Protocol
SR	Super Resolution
SSH	Secure SHell
SVM	Support Vector Machine
TM	Tele Management
TMF	Tele Management Forum
TPR	True Positive Rate
UNSW	University of New South Wales
URLLC	Ultra Reliable Low Latency Communications
VAE	Variational Autoencoder
VLAN	Virtual Local Area Network
VPN	Virtual Private Network
WAN	Wide Area Network
WGAN	Wasserstein Generative Adversarial Network
WMI	Windows Management Instrumentation
XML	Extensible Markup Language
YAML	YAML Ain't Markup Language
ZSM	Zero-touch network and Service Management

References

[1] Dmytriiev, O., 2025. ‘Orchestrating the future of fully autonomous networks with GenAI’. [online] *Nokia*. Available at: https://www.nokia.com/blog/orchestrating-the-future-of-fully-autonomous-networks-with-genai. Accessed 28 Aug. 2025.

[2] Tavarez, G., 2024. ‘NetBrain R12: Proactive Network Management with AI and Automation’. [online] *Cloud Computing TMCnet*. Available at: https://cloud-computing.tmcnet.com/breaking-news/articles/461209-netbrain-r12-proactive-network-management-with-ai-automation.htm. Accessed 28 Aug. 2025.

[3] Liu, Y., Du, H., Niyato, D., Kang, J., et al., 2023. ‘Deep Generative Model and Its Applications in Efficient Wireless Network Management: A Tutorial and Case Study’. *arXiv preprint* arXiv:2303.17114. Available at: https://doi.org/10.48550/arxiv.2303.17114.

[4] Celik, A., Abdallah, A., Alkhateeb, A. and Eltawil, A.M., 2024. ‘At the Dawn of Generative AI Era: Generative Models and New Frontiers in 6G Wireless Intelligence’. *IEEE Future Networks World Forum*. Available at: https://fnwf2024.ieee.org/program/tutorials/dawn-generative-ai-era-generative-models-and-new-frontiers-6g-wireless. Accessed 28 Aug. 2025.

[5] [x]cube LABS, 2024. ‘Exploring Zero-Shot and Few-Shot Learning in Generative AI’. [online] *[x]cube LABS*. Available at: https://www.xcubelabs.com/blog/exploring-zero-shot-and-few-shot-learning-in-generative-ai/. Accessed 28 Aug. 2025.

[6] Huo, P., Sridhar, A.N., Khan, M.F.F., Maeng, K., et al. ‘QoS-Diff: Adaptive Auto-Tuning Framework for Low-Latency Diffusion Model Inference’. 3 Dec. 2024, pp. 1–7, https://doi.org/10.1145/3696409.3700277. Accessed 28 Aug. 2025.

[7] Rao, Y.N. and Babu, K.S. (2023) ‘An imbalanced generative adversarial network-based approach for network intrusion detection in an imbalanced dataset’, *Sensors*, 23(1), p.550. https://doi.org/10.3390/s23010550.

[8] Chen, Y., Xie, H., et al. (2023) ‘Automatic Root Cause Analysis via Large Language Models for Cloud Incidents’, *arXiv preprint*, arXiv:2305.15778. https://doi.org/10.48550/arxiv.2305.15778.

[9] Mistry, H. K., Mavani, C., Goswami, A., and Patel, R. ‘Artificial intelligence for networking.’ *Educational Administration: Theory and Practice*, 30, no. 7 (2024): 813-821. https://kuey.net/index.php/kuey/article/view/6854.

[10] Bovenzi, G., Cerasuolo, F., Ciuonzo, D., Di Monda, D., et al. (2025) ‘Mapping the Landscape of Generative AI in Network Monitoring and Management’, *arXiv preprint*, arXiv:2502.08576. https://doi.org/10.48550/arXiv.2502.08576.

[11] Mani, S.K., Zhou, Y., Hsieh, K., Segarra, S., et al., 2023. ‘Enhancing network management using code generated by large language models’. *Proceedings of the ACM SIGCOMM 2023 Workshop on Network-Application Integration (NAI '23)*, pp.1–7. Available at: https://doi.org/10.1145/3626111.3628183.

[12] Navidan, H., Fard Moshiri, P., Nabati, M., Shahbazian, R., et al., 2021. ‘Generative Adversarial Networks (GANs) in networking: A comprehensive survey & evaluation’. *Computer Networks*, 194, p.108149. Available at: https://doi.org/10.1016/j.comnet.2021.108149.

[13] Berahmand, K., Daneshfar, F., Salehi, E.S., Li, Y. & Xu, Y., 2024. ‘Autoencoders and their applications in machine learning: a survey’. *Artificial Intelligence Review*, 57, article 28. Available at: https://doi.org/10.1007/s10462-023-10662-6.

[14] Liang, R., Yang, B., Chen, P., Li, X., et al., 2025. ‘Diffusion models as network optimizers: explorations and analysis’. *IEEE Internet of Things Journal*, 12(10), pp.13183–13193. Available at: https://doi.org/10.1109/JIOT.2025.3528955.

[15] Tang, F., Wang, X., Yuan, X., Luo, L., et al., 2025. ‘Large language model (LLM) assisted end-to-end network health management based on multi-scale semanticization’. *arXiv preprint* arXiv:2406.08305. Available at: https://doi.org/10.48550/arXiv.2406.08305.

[16] Muhammad, K., David, T., Nassisid, G. & Farus, T., 2025. ‘Integrating generative AI with network digital twins for enhanced network operations’. *arXiv preprint* arXiv:2406.17112. Available at: https://doi.org/10.48550/arXiv.2406.17112.

[17] Huang, X., Yang, H., Zhou, C., He, M., Shen, X. & Zhuang, W., 2024. ‘When digital twin meets generative AI: intelligent closed-loop network management’. *arXiv preprint* arXiv:2404.03025. Available at: https://doi.org/10.48550/arXiv.2404.03025.

[18] Tan, T., Tang, F., Luo, L., Wang, X., Li, Z. & Zhao, M., 2025. ‘Adapting network information into semantics for generalizable and plug-and-play multi-scenario network diagnosis’. *arXiv preprint* arXiv:2501.16842. Available at: https://doi.org/10.48550/arXiv.2501.16842.

[19] G. Hollósi, D. Ficzere and P. Varga, ‘Generative AI for Low-Level NETCONF Configuration in Network Management Based on YANG Models.’ *2024 20th International Conference on Network and Service Management (CNSM)*, Prague, Czech Republic, 2024, pp. 1-7. Available at: https://opendl.ifip-tc6.org/db/conf/cnsm/cnsm2024/1571045554.pdf.

[20] S. D'Urso, B. Martini and F. Sciarrone, ‘A Novel LLM Architecture for Intelligent System Configuration,’ *2024 28th International Conference Information Visualisation (IV)*, Coimbra, Portugal, 2024, pp. 326-331. Available at: https://www.proceedings.com/content/076/076965webtoc.pdf.

[21] Wang, C., Scazzariello, M., Farshin, A. & Ferlin, S., 2024. ‘NetConfEval: Can LLMs facilitate network configuration?’ *Proceedings of the ACM on Networking*, 2(CoNEXT2), article 7. Available at: https://doi.org/10.1145/3656296.

[22] Lira, O.G., Caicedo, O.M. & da Fonseca, N.L.S., 2024. ‘Large language models for zero touch network configuration management’. *arXiv preprint* arXiv:2408.13298. Available at: https://doi.org/10.48550/arXiv.2408.13298.

[23] D. Brodimas et al., ‘Towards Intent-based Network Management for the 6G System adopting Multimodal Generative AI,’ *2024 Joint European Conference on Networks and Communications & 6G Summit (EuCNC/6G Summit)*, Antwerp, Belgium, 2024, pp. 848-853. Available at: https://zenodo.org/records/11210729.

[24] Patras5G OpenSlice – a prototype open-source and standards-based Operations Support System (OSS) for delivering Network as-a-Service (NaaS) – https://patras5g.eu/.

[25] S. Piroti, A. Chawla and T. Zanouda, ‘Mobile Network Configuration Recommendation Using Deep Generative Graph Neural Network’, *IEEE Networking Letters*, vol. 6, no. 3, pp. 179-182, Sept. 2024. Available at: https://arxiv.org/abs/2406.04779.

[26] Wang, C., Scazzariello, M., Farshin, A., Kostic, D. & Chiesa, M., 2023. ‘Making network configuration human friendly’. *arXiv preprint* arXiv:2309.06342. https://doi.org/10.48550/arXiv.2309.06342.

[27] Y. Fang, K. Lu, J. Xue, F. Li and Z. Lyu, ‘LLMNDC: A Novel Approach for Network Device Configuration based on Fine-tuned Large Language Models’, *2024 5th International Conference on Computer Engineering and Intelligent Control (ICCEIC)*, Guangzhou, China, 2024, pp. 283-289. Available at: https://www.researchgate.net/publication/388422575_LLM4DistReconfig_A_Fine-tuned_Large_Language_Model_for_Power_Distribution_Network_Reconfiguration.

[28] S. Chakraborty, N. Chitta and R. Sundaresan, ‘Automation of Network Configuration Generation using Large Language Models’, *2024 20th International Conference on Network and Service Management (CNSM)*, Prague, Czech Republic, 2024, pp. 1-7. https://www.researchgate.net/publication/387590920_Automation_of_Network_Configuration_Generation_using_Large_Language_Models.

[29] D. Donadel, F. Marchiori, L. Pajola and M. Conti, ‘Can LLMs Understand Computer Networks? Towards a Virtual System Administrator’, *2024 IEEE 49th Conference on Local Computer Networks (LCN)*, Normandy, France, 2024, pp. 1-10. https://arxiv.org/pdf/2404.12689.

[30] Mondal, R., Tang, A., Beckett, R., Millstein, T. & Varghese, G., 2023. ‘What do LLMs need to synthesize correct router configurations?’ *Proceedings of the 22nd ACM Workshop on Hot Topics in Networks (HotNets ’23)*, pp.1–7. Available at: https://doi.org/10.1145/3626111.3628194.

[31] E. -D. Jeong, H. -G. Kim, S. Nam, J. -H. Yoo, et al., 'S-Witch: Switch Configuration Assistant with LLM and Prompt Engineering,' *NOMS 2024-2024 IEEE Network Operations and Management Symposium*, Seoul, Korea, Republic of, 2024, pp. 1-7. https://www.researchgate.net/publication/381935952_S-Witch_Switch_Configuration_Assistant_with_LLM_and_Prompt_Engineering.

[32] Tang, T., Yao, J., Wang, Y., Sha, Q., et al. (2025, April). 'Application of Deep Generative Models for Anomaly Detection in Complex Financial Transactions'. *2025 4th International Conference on Artificial Intelligence, Internet and Digital Economy (ICAID)* (pp. 133-137). IEEE. https://arxiv.org/abs/2504.15491.

[33] Ericson, L., Zhu, X., Han, X., Fu, R., et al., 2024. 'Deep generative modeling for financial time series with application in VaR: a comparative review'. *arXiv preprint.* https://doi.org/10.48550/arXiv.2401.10370.

[34] Lee, D.K.C., Guan, C., Yu, Y. & Ding, Q., 2024. 'A comprehensive review of generative AI in finance'. *FinTech*, 3(3), pp.460–478. https://doi.org/10.3390/fintech3030025.

[35] Saxena, A., Mahajan, J., & Verma, S. (2024). Generative AI in Banking Financial Services and Insurance. Springer. Available at: https://content.e-bookshelf.de/media/reading/L-24431748-b03114f707.pdf.

[36] Liu, Y., Du, H., Niyato, D., Kang, J., et al., 2024. 'Deep generative model and its applications in efficient wireless network management: a tutorial and case study'. *IEEE Wireless Communications*, 31(4), pp.199–207. Available at: https://doi.org/10.1109/MWC.009.2300165.

[37] Xiao, S., He, D. & Gong, Z., 2018. 'Deep-Q: Traffic-driven QoS inference using deep generative network'. *NetAI '18: Workshop on Network Meets AI & ML*, pp.67–73. Available at: https://doi.org/10.1145/3229543.3229549.

[38] V. Gudepu, B. Chirumamilla, V. Reddy Chintapalli, P. Castoldi, L. Valcarenghi and K. Kondepu, 'Generative Adversarial Networks-Based AI/ML Model Adaptive Retraining for Beyond 5G Networks', *European Wireless 2023; 28th European Wireless Conference*, Rome, Italy, 2023, pp. 224-229. https://arxiv.org/abs/2408.14827v1.

[39] Hua, Y., Li, R., Zhao, Z., Chen, X. & Zhang, H., 2020. 'GAN-powered deep distributional reinforcement learning for resource management in network slicing'. *IEEE Journal on Selected Areas in Communications*, 38(2), pp.334–349. Available at: https://doi.org/10.1109/JSAC.2019.2959185.

[40] Fan, J., Wang, Z., Xie, Y., & Yang, Z. (2020, July). 'A theoretical analysis of deep Q-learning'. *Learning for dynamics and control* (pp. 486-489). PMLR. https://arxiv.org/abs/1901.00137.

[41] Li, S., Lin, X., Liu, Y., Li, G., et al. (2024) 'OpticGAI: Generative AI-aided Deep Reinforcement Learning for Optical Networks Optimization', *Proceedings of the 1st SIGCOMM Workshop on Hot Topics in Optical Technologies and Applications in Networking (HotOptics '24)*. Association for Computing Machinery, New York, NY, USA, pp. 1–6. https://arxiv.org/abs/2406.15906.

[42] Park, C., Lee, J., Kim, Y., Park, J.-G., et al. (2022). 'An enhanced AI-based network intrusion detection system using generative adversarial networks'. *IEEE Internet of Things Journal*. Available at: https://ieeexplore.ieee.org/document/9908159.

[43] García González, G., Casas, P., Fernández, A. and Gómez, G. (2020) 'On the Usage of Generative Models for Network Anomaly Detection in Multivariate Time-Series', *arXiv preprint*. https://arxiv.org/abs/2010.08286.

[44] Sindiramutty, S. R., Prabagaran, K. R., Jhanjhi, N. Z., Murugesan, et al. (2025). 'Generative AI in Network Security and Intrusion Detection'. *Reshaping CyberSecurity With Generative AI Techniques* (pp. 77-124). IGI Global Scientific Publishing. https://doi.org/10.4018/979-8-3693-5415-5.ch003.

[45] Sinha, H. (2024). 'The Identification of Network Intrusions with Generative Artificial Intelligence Approach for Cybersecurity'. *Journal of Web Applications and Cyber Securit*y (e-ISSN: 2584-0908), 2(2), 20–29. https://doi.org/10.48001/jowacs.2024.2220-29.

[46] García González, G., Casas, P., Fernández, A. and Gómez, G. (2021) 'On the Usage of Generative Models for Network Anomaly Detection in Multivariate Time-Series', *SIGMETRICS Performance Evaluation Review*, 48(4), pp. 49–52. https://arxiv.org/abs/2010.08286.

[47] Li, D., Chen, D., Jin, B., Shi, L., Goh, J. and Ng, S.K. (2019) 'MAD-GAN: Multivariate Anomaly Detection for Time Series Data with Generative Adversarial Networks'. *Artificial Neural Networks and Machine Learning – ICANN 2019: Text and Time Series*. Lecture Notes in Computer Science, vol. 11730. https://doi.org/10.1007/978-3-030-30490-4_56.

[48] Poongodi, M., & Hamdi, M. (2023). ‘Intrusion detection system using distributed multilevel discriminator in GAN for IoT system’. *Transactions on Emerging Telecommunications Technologies*, 34(11), e4815. https://doi.org/10.1002/ett.4815.

[49] MSSD-MCMC/sigma-rules: Updated Suricata rules specifically tailored for Project SIGMA, aimed at enhancing network threat detection – https://github.com/MSSD-MCMC/sigma-rules

[50] Suricata is an open-source network threat detection engine developed by the Open Information Security Foundation – https://suricata.io/

[51] Cisco AI Network Analytics – https://www.cisco.com/site/us/en/solutions/artificial-intelligence/ai-assistant/index.html

[52] Juniper Mist AI – https://www.juniper.net/gb/en/mist-ai-native-networking-platform.htm

[53] Fortinet (2024) *Fortinet makes network operations simpler with generative AI*. Available at: https://www.fortinet.com/blog/business-and-technology/fortinet-makes-network-operations-simpler-with-generative-ai

[54] IBM Watsonx – https://www.ibm.com/us-us/watsonx

[55] NetBrain’s Chatbot – https://www.netbraintech.com/product/chatbot/

[56] Atera's AI-Powered IT Management Platform – https://www.atera.com/

[57] ServiceNow IT Infrastructure Management – https://www.servicenow.com/lpdem/demonow-it-infrastructure.html

[58] OpenAI Codex – https://openai.com/codex/

[59] Broadcom DX Operational Observability – https://techdocs.broadcom.com/us/en/ca-enterprise-software/it-operations-management/dx-operational-observability/saas.html

[60] XenonStack GenAI for NOC – https://www.xenonstack.com/blog/generative-ai-for-network-operations

[61] NetGPT: A GPT-based AI tool to automate network engineering tasks – https://github.com/KennethGrace/NetGPT

[62] LangChain – https://github.com/langchain-ai/langchain

[63] Karapantelakis, A., Thakur, M., Nikou, A., Moradi, F., Olrog, C. and Gaim, F. (2024) ‘Using Large Language Models to Understand Telecom Standards’, in *Proceedings of the 2024 IEEE International Conference on Machine Learning and Communication Networks (ICMLCN)*. IEEE, pp. 1–6. https://doi.org/10.1109/ICMLCN59089.2024.10624786

[64] Abane, A., Battou, A. and Merzouki, M. (2024) ‘An adaptable AI assistant for network management’, in *Proceedings of NOMS 2024 – IEEE Network Operations and Management Symposium*. IEEE, Seoul, Korea, pp. 1–6. https://doi.org/10.1109/NOMS59830.2024.10574957

[65] Pydantic GenAI Python framework – https://ai.pydantic.dev/

[66] Google’s Agent Development Kit (ADK) – https://google.github.io/adk-docs/

[67] Microsoft AutoGen – https://www.microsoft.com/en-us/research/project/autogen/

[68] AI/ML in Cisco Catalyst Center – https://www.ciscolive.com/c/dam/r/ciscolive/emea/docs/2025/pdf/IBOOPS-2391.pdf

[69] Cisco AI Assistant – https://www.cisco.com/site/us/en/solutions/artificial-intelligence/ai-assistant/index.html

[70] Juniper Mist AI – https://www.juniper.net/gb/en/mist-ai-native-networking-platform.html

[71] Juniper Marvis AI Assistant – https://www.juniper.net/content/dam/www/assets/datasheets/us/en/cloud-services/marvis-ai-assistant-datasheet.pdf

[72] Krékity, G. (2025) *Juniper Mist AI: Revolution in the Network*. https://socwise.eu/juniper-mist-ai-revolution-in-the-network/

[73] Liu, Y., Jia, Y., Geng, R., Jia, J. and Gong, N.Z. (2024) ‘Formalizing and Benchmarking Prompt Injection Attacks and Defenses’, in *Proceedings of the 33rd USENIX Security Symposium (USENIX Security 24)*. USENIX Association, pp. 1–14. https://arxiv.org/abs/2310.12815.

[74] Freitas, S., Kalajdjieski, J., Gharib, A. & McCann, R., 2024. ‘AI-Driven Guided Response for Security Operation Centers with Microsoft Copilot for Security’. *arXiv preprint* arXiv:2407.09017. Available at: https://doi.org/10.48550/arXiv.2407.09017

[75] Charlotte AI – https://www.crowdstrike.com/en-us/platform/charlotte-ai/

[76] Darktrace ActiveAI Security Platform – https://www.darktrace.com

[77] Zscaler Zero Trust Exchange – https://www.zscaler.com/

[78] SentinelOne Purple AI – https://www.sentinelone.com/

[79] Tenable ExposureAI – https://www.tenable.com

[80] Coralogix – https://coralogix.com/

[81] Lakera Guard – https://www.lakera.ai/lakera-guard

[82] Calypso AI – https://calypsoai.com/

[83] Lasso Security – https://www.lasso.security

[84] LlamaFirewall – https://meta-llama.github.io/PurpleLlama/LlamaFirewall/

[85] Adversarial Robustness Toolbox (ART) – https://github.com/Trusted-AI/adversarial-robustness-toolbox?tab=readme-ov-file

[86] Garak – https://github.com/NVIDIA/garak?tab=readme-ov-file

[87] Zeek – https://docs.zeek.org/en/current/about.html

[88] Zeek-MCP – https://mcpmarket.com/server/zeek

[89] Cybersecurity AI framework for AI Security – https://github.com/aliasrobotics/cai?tab=readme-ov-file#-milestones

[90] Mayoral-Vilches, V., Navarrete-Lozano, L.J., et al. 2025. ‘CAI: An Open, Bug Bounty-Ready Cybersecurity AI’. *arXiv preprint* arXiv:2504.06017. https://doi.org/10.48550/arXiv.2504.06017