International Journal of Emerging Research in Engineering and Technology
Pearl Blue Research Group | Volume 7 Issue 2 PP 203-209, 2026
ISSN: 3050-922X | https://doi.org/10.63282/3050-922X.IJERET-V7I2P125

Original Article

A Systematic Survey of Autonomous AIOps and Generative AI in Cloud-Native Infrastructure

Dr. Prashant Kumar Srivastava[1], Sanjeev Agrawal[2]
[1]Associate Professor, SOCT.
[2]Global Educational (SAGE) University, Bhopal.

Received On: 02/04/2026 ***Revised On: 01/05/2026*** ***Accepted On: 09/05/2026*** ***Published On: 15/05/2026***

Abstract *- Cloud-native infrastructures have transformed modern computing through scalable, flexible, and resilient deployment models based on microservices, containers, orchestration, and DevOps practices. Nevertheless, increasing system complexity, distributed workloads, and dynamic operational environments create major challenges in monitoring, fault management, security, and resource optimization. This systematic survey reviews the integration of Autonomous Artificial Intelligence for IT Operations (AIOps) and Generative Artificial Intelligence within cloud-native infrastructure. The study examines cloud-native architectural foundations and explores how artificial intelligence and machine learning enhance operational automation and infrastructure intelligence. Core AIOps capabilities, including automated monitoring, anomaly detection, predictive maintenance, and root cause analysis, are analyzed alongside Generative AI applications in log analysis, incident management, and decision support. Recent literature demonstrates that combining AIOps with Generative AI improves reliability, reduces mean time to recovery, optimizes resource utilization, and minimizes manual intervention. The survey also identifies key challenges, including model drift, explainability limitations, privacy concerns, integration complexity, security risks, and computational demands. Finally, future research directions toward secure, scalable, trustworthy, and self-healing autonomous cloud operations are highlighted for future next-generation digital infrastructure ecosystems.*

Keywords *- Autonomous AIOps, Generative AI, Cloud-Native Infrastructure, Anomaly Detection, Intelligent Cloud Operations.*

1. Introduction

Cloud computing and cloud-native architectures have revolutionized IT infrastructure by providing scalable, flexible computing resources available on demand[1]. The cloud-native architecture uses microservices, containers, orchestration, and DevOps to provide fast application deployment and better resource utilization[2]. Nevertheless, the complexity, workload dynamics, and distributed nature of cloud infrastructures pose many challenges to monitoring, fault management, performance optimization, and security. Such challenges have resulted in increased demand for intelligent automation techniques such as AIOps and Generative AI for cloud infrastructure[3]. The term "AIOps" refers to systems that use AI and big data to improve and automate IT operations, including event correlation, anomaly detection, root cause investigation, and remediation. AIOps is a system that links observability, recognizing continuous events with comprehension, interpreting their relevance and deciding appropriate action so transcending simple buzzword status. AIOps systems can continuously ingest and learn patterns in operational data that humans might not pick up on [4].

AI for IT Operations, or AIOps, is the field of research dedicated to using AI for IT Services management and improvement. The goal of analytics, big data, and machine learning is to monitor computer infrastructure and offer proactive insights and recommendations to increase MTTR, distribute computing resources efficiently, and reduce failures [5]. From simple scheduling and resource management to advanced failure management activities like anomaly detection, remediation, and failure prediction, AIOps has a vast and varied range of tools for all of these uses. However, as a relatively new, cross-disciplinary field, AIOps remains largely unstructured as a research area.

Generative AI is a transformative form of AI that generates new content, including text, images, code and analytical outputs, by learning patterns from large-scale datasets[6]. Unlike traditional AI systems that primarily focus on prediction and classification tasks, Generative AI creates new and context-aware responses using advanced deep learning models. Foundation models and large language models (LLMs), especially transformer-based architectures, have greatly improved Generative AI capabilities across a wide range of domains [7]. These models can interpret complex instructions, reason, summarize information and support automated decision-making processes. Generative AI is increasingly used in today’s IT and cloud environments to enhance operational intelligence, automate repetitive tasks, and improve human–machine collaboration.

1.1. Structure of the Paper

This paper is organized as follows: **Section II** discusses cloud-native infrastructure and AI integration. **Section III** presents Autonomous AIOps and its key technologies. **Section IV** explores Generative AI in cloud operations and related challenges. **Section V** reviews existing literature and comparative studies. Finally, **Section VI** concludes the paper and highlights future research directions.

2. Cloud-Native Infrastructure and AI Integration

Modern application development increasingly relies on distributed architectures that emphasize agility, scalability, and resilience. This approach utilizes microservices, containers, orchestration platforms, and automated deployment pipelines to support rapid application delivery and efficient workload management. Compared with traditional monolithic systems, these environments provide greater flexibility and reliability while enabling organizations to adapt quickly to changing demands. Furthermore, intelligent technologies enhance operational efficiency by streamlining processes and making decisions based on data. ML and predictive analytics can assist in the monitoring of performance, detecting anomalies, optimizing resources, and anticipating and managing issues, all of which help create more autonomous and efficient computing environments [8].

2.1. Cloud-Native Infrastructure Design and Architecture

The use of modern cloud-based technologies and distributed computing concepts is becoming more and more crucial to scalable, resilient and flexible application development. These environments are based on microservices architecture, as opposed to the monolithic system architecture, which breaks applications into loosely coupled, independent services, allowing for quicker development, deployment and maintenance. Resource-efficient use, resource auto-scaling and failover further enhance application reliability and support in dynamic cloud environments with the help of containerization technologies and orchestration platforms. Further, DevOps practices and continuous integration/continuous deployment (CI/CD) pipelines aid the speed of deployment and ensure agility in operation. Cloud-native infrastructure's architectural features are well-suited for current digital services and large-scale distributed systems [9].

2.2. Core Technologies Enabling Cloud-Native Systems

Cloud-native means that modern technologies are integrated that work together to create scalable, flexible and highly automated applications in the distributed cloud. Every technology serves a specific purpose in improving performance, reliability and manageability.

- Containerization (Docker): Provides lightweight application environments that can run easily and can be moved from one place to another, which leads to consistency between development, test and production environments.
- Microservices Architecture: Decomposes applications into small autonomous services, increases agility and fault isolation [10].
- DevOps and CI/CD Pipelines: Automates software development lifecycle processes, such as building[11], testing, and deploying, enabling continuous delivery and faster release cycles.
- Service Mesh (Istio, Linkerd): Provides secure, reliable communication between microservices, with features such as traffic management, encryption, and observability.
- Observability Tools (Monitoring, Logging, Tracing): Streamlines the process of detecting and fixing errors by collecting relevant system data and logs in real-time.
- Cloud Infrastructure Services (IaaS, PaaS, SaaS): Provides scalable compute, storage, and networking resources that form the foundation for cloud-native application deployment.

2.3. Role of AI and ML in Optimizing Cloud-Native Operations

Cloud-native systems are undergoing a revolutionary shift due to AI and ML, which are automating decision-making and allowing predictive capabilities. By using ML algorithms combined with cloud-native orchestration platforms, such as Kubernetes, one can leverage usage analysis data to effectively distribute resources and optimize costs. Thus, it becomes possible to achieve dynamic scalability for variable systems, such as e-commerce sites[12]. Observability tools based on AI technologies, such as Dynatrace and New Relic, can process massive amounts of telemetry data to discover anomalies, predict failures and generate insights. These tools enable one to resolve performance issues, such as latency, by tracing their sources in advance. Furthermore, AI and ML technologies can be used to detect abnormal patterns in the behaviour of cloud-native systems and their networking environment [13][14].

3. Overview of Autonomous Aiops

Autonomous AIOps refers to the integration of AI, ML, and automation technologies to manage and optimize IT operations with minimal human intervention. It enables a shift from traditional reactive management to proactive and self-healing operational environments by continuously monitoring infrastructure, analyzing large volumes of operational data, and automatically responding to potential issues.

3.1. Autonomous AIOps Architecture and Core Components

AIOps is the application of artificial intelligence to enhance and automate IT operations through intelligent data processing and decision-making. This paradigm shifts from a reactive to a proactive approach to IT management and enables systems to anticipate, detect, and correct problems automatically, while maximizing the utilization of resources within a complex cloud environment. The AIOps architecture provides a complex structure that starts from data ingestion and integration layers, which constantly gather and normalize data from various cloud infrastructure sources [15]. This information is fed into the real-time operation of advanced analytics engines, which leverage ML algorithms to generate insights and forecasts on system behaviour and potential problems. These insights are used in automation frameworks to perform remediation actions that are learned from patterns and policies. It includes visualization and reporting features that transform complex operational data into actionable intelligence, enabling operators to make informed choices and stay informed about automated processes. The integration enables components to communicate seamlessly and offers the flexibility to accommodate changing operational demands.

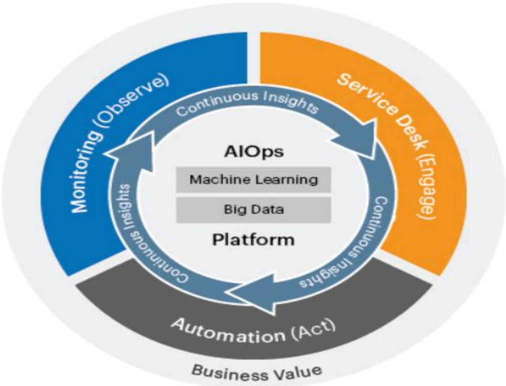

Fig 1: AIOps Architecture and Operational Framework

The AIOps framework integrates monitoring, human service desk interactions, and automation, while leveraging machine learning and big data analytics to generate more value and efficiency from IT operations. (See Fig. 1).

3.1. Automated Monitoring and Anomaly Detection

One of the most critical capabilities of AIOps is the ability to analyze and continuously monitor IT environments. In large-scale cloud and distributed systems with dynamic workloads and vast amounts of data, traditional monitoring techniques based on static thresholds and rule-based alerts are less effective [16]. To address these challenges, AIOps platforms collect and process operational information from sources like logs, metrics, traces and events from different infrastructure and application resources. They can be correlated on the fly via ML/BD and abnormal behaviour/operational issues can be automatically detected [17][18].

Failure, degradation and potential threats to security can be caught early by a deviation from normal system behavior. Machine learning models offer more accurate detection, decrease false alarms and capture users' normal patterns of activities without the need for complex rules or thresholds [19]. AIOps platforms can bridge data sources to provide context, facilitate root cause analysis, and accelerate incident resolution. These features provide proactive and self-healing IT operations, helping to ensure system reliability and efficiency [20].

3.2. Key Technologies and Techniques in AIOps

The implementation of AIOps involves leveraging AI, ML, and other advanced techniques in data analysis to address the increasing complexity of IT environments. These technologies enable the collection, processing, and analysis of enormous amounts of information produced by infrastructure, applications, networks and services while they are operating. This data can uncover patterns and relationships that can be utilized for intelligent decisions, system automation, and proactive system management. Several core techniques of AIOps, including machine learning algorithms, anomaly detection, predictive maintenance, and automated root cause analysis, contribute to improved system reliability, performance, and operational efficiency [21].

- Machine Learning Algorithms:Machine learning algorithms play a significant role in enabling intelligent data analysis for AIOps, as they process large volumes of operational data to identify patterns, gain insights, and support automated decision-making.
- Supervised Learning: The models created from the labelled datasets are widely employed in classification and prediction tasks, enabling the prediction of incidents, service failures, or performance-related issues with a high degree of accuracy.
- Unsupervised Learning: The methods of pattern discovery can analyze data without any prior labels or descriptions and uncover structures, correlations and unusual behaviours that can be useful in environments that are dynamic and complex within the IT system.
- Anomaly Detection: The system constantly observes activities and employs advanced statistical and machine learning technologies to identify patterns that deviate from the normal, thus allowing the early detection of faults, degradation of operation, and potential security threats.
- Predictive Maintenance: The history of operation and predictive models are used to estimate the chance of failure of mechanical or electrical components or programs and maintenance can be planned in advance of a major failure.
- Downtime Reduction and Asset Optimization: Proactive maintenance streamlines to lower operational costs, prolong life of IT assets, components of infrastructure, and contributes to lower service disruptions.
- Automated Root Cause Analysis (RCA): Patterns of incidents are automatically determined using correlation and analytical techniques, to help to diagnose incidents faster, reducing troubleshooting and enabling more rapid service recovery.

4. Generative AI for Intelligent IT and Cloud Operations

Today's IT environments and cloud services are becoming more complex, and require intelligent automation, adaptive decision-making, and effective management of operations. In recent years, AI has emerged with new functionalities, such as generative AI, which goes beyond mere prediction and classification. Among these developments are the advanced learning capabilities of Generative AI, which can create human-like text, insights, operational summaries, and context-aware recommendations.

These capabilities enable the analysis of vast and vast quantities of logs, metrics, and system events in cloud and IT operations, for detecting anomalies, determining root causes and initiating automated incident responses. Additionally, the ability of AIOps platforms to integrate with LLMs and generative technologies enhances the autonomous and adaptive nature of IT management, thereby minimizing the need for manual processes and enhancing system reliability, efficiency, and service uptime.

4.1. Foundations of Generative AI in IT Operations

Generative AI is a very advanced branch of AI that leverages patterns learned from vast amounts of data to produce new content like text, code, or analytics. While the primary function of conventional AI is prediction and classification, generative AI provides contextual responses using deep learning and transformer architectures. The emergence of foundation models and large language models (LLMs) has greatly improved their linguistic understanding capabilities. Generative AI improves IT processes through intelligent capabilities such as log analytics, incident management, automation, and decision-making support. By analyzing data from operational processes and identifying useful patterns, Generative AI helps IT professionals understand how systems operate and solve issues more effectively. Furthermore, generative AI provides automated scripting, troubleshooting, and other recommendations that enhance intelligence in IT operations and reduce effort. The major areas of functionality for Generative AI in IT operations are shown in Figure 2 below.

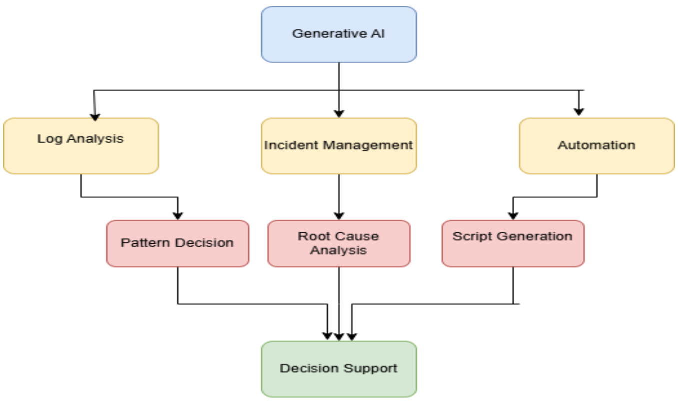

Fig 2: Major Functional Areas of Generative AI in IT Operations

4.1. GenAI for Log Analysis and Root Cause Analysis

Generative AI has emerged as an effective approach for log analysis and root cause analysis (RCA) in modern IT operations. Modern cloud-native and distributed systems produce large volumes of unstructured logs and operational events that are time-consuming to analyze manually and error-prone. Transformer models and LLMs can analyze and understand log data, detect anomalies, and enable automatic fault diagnosis. By leveraging generative AI, one can discover relevant patterns in logs, correlate operational events and summarize operational information to speed up problem investigation. Moreover, the use of LLMs enables the identification of possible root causes of incidents by analyzing logs, metrics and previously recorded incident data[22]. Consequently, Generative AI contributes to improving operational transparency and facilitating intelligent IT operations.

4.2. Challenges and Limitations of Generative AI in Cloud Operations

Integrating Generative AI with cloud automation has many advantages, but there are also major hurdles that need to be overcome before Generative AI components can be deployed and maintained successfully. The efficacy of cloud-based AI systems may be compromised by these technological and operational concerns [23].

- Model Drift**:** AI models are constantly changing, and they can lose accuracy over time and need to be retrained and monitored periodically in cloud environments.
- Data Privacy and Compliance: Generative AI uses logs and operational data that may contain sensitive or personally identifiable data (PII), raising privacy and regulatory concerns.
- Integration Complexity**:** Integrating Generative AI into current systems and cloud-based solutions may have compatibility and architecture challenges.
- Explainability and Transparency**:** The decisions made by generative AI models are often described as “black-box” decisions, which can be challenging to interpret during critical incidents.
- Security Risks**:** Adversarial attacks and tampered inputs into AI systems can compromise cloud security and the trustworthiness of automation.
- Resource and Computational Constraints**:** Training and deployment of big AI models is expensive and time-consuming.
- Ethical and Governance Concerns**:** There are questions about accountability, human oversight, and ethics behind decisions made without human intervention when using AI-driven automation.

5. Literature Review

This section highlights recent trends and findings in the field of AIOps and Generative AI within cloud and cloud-native landscapes, covering advancements in predictive maintenance, autonomous incident management, resource optimization, and operational automation to improve system reliability, scalability, and overall efficiency.

H. N. Gadbail et al. (2026) discuss many aspects of Cloud Computing, including service models, virtualization, fault tolerance, business models and user experience. The article also discusses how AI technology can enhance Cloud environments through automation, predictive analytics, improved data management, and tools such as Chatbot and AIOps. The combination of AI and Cloud Computing increases productivity, reliability, and decision-making accuracy and supports advanced workloads that businesses cannot support on their own[24].

V. Raj (2025) indicated that there were various weaknesses in the traditional reactive maintenance approaches, which were characterized by unpredictable downtimes and interruptions in the performance of operations. To address these problems, the authors analyzed the adoption of AIOps for operations management in a hybrid cloud computing environment. As a result of using the technology, operations could be predicted and potential problems avoided before they occurred [25].

D. K. Seth, K. K. Ratra, and A. P. Sundareswaran (2025) addressed the application of AI and Generative AI (GenAI) for the automation and optimization of cloud and hybrid cloud infrastructures. Workload management and system performance improved through AI-powered resource scaling, while security and compliance automation enhanced protection and reliability. And GenAI is assisting with cloud orchestration and operational workflows. The results demonstrated lower operational costs, quicker service deployment and better efficiency in managing cloud infrastructure [26].

W. C. Potts and C. Carver (2024) explore AIOps, which was studied based on its suitability in both governmental and enterprise settings. It was important to emphasize the interconnectivity among various technologies, as well as their interdependencies, when discussing challenges in the adoption and implementation of AIOps, and the possibility of consolidating technology silos by combining the power of data science, IT operations, and software development. It was found that AIOps' effectiveness largely depended not only on the availability of specific tools but also on organizational capabilities, communication, process engineering, and alignment with business priorities [27].

P. K. Thota (2024) proposes a Generative AI-based framework for autonomous cloud infrastructure management. The framework used real-time decision-making, serverless computing, event-driven workflows, and reinforcement learning for cloud operations automation. AI-driven automation with APIs and IaC to optimize resource allocation, predict failures and manage security risks. The results showed improved scalability, lower operational costs, improved system availability and efficient cloud resource management [28].

S. Sundar Ray(2023) proposed a Generative AI-based autonomous incident response framework for distributed and cloud-native enterprise systems. The framework used retrieval-augmented generation (RAG), vector embeddings and agentic remediation workflows to allow automated root cause analysis and intelligent remediation. Experimental results demonstrated the effectiveness of Generative AI in autonomous cloud operations by showing significant improvements in operational efficiency, such as lower mean time to resolution (MTTR) and reduced manual debugging effort [29].

A. Hayes et al. (2022) propose a generative AI-based framework for automated log summarization and root-cause triage of distributed systems. They applied a framework that combined large language models, retrieval-augmented generation, and evidence mapping to generate context-aware summaries of unstructured logs and detect anomalies for faster diagnostics. The results demonstrated the potential of Generative AI for intelligent and scalable cloud operations, showing reduced incident triage time, lower mean time to resolution, and improved operational efficiency [30].

Table I summarizes recent studies on Autonomous AIOps and Generative AI, highlighting advancements in intelligent automation, predictive analytics, incident management, and cloud optimization, while identifying challenges and future research opportunities.

Table 1: Comparative Analysis of Autonomous Aiops and Generative Ai in Cloud-Native

Authors	Focus Area	Techniques	Key Findings	Limitations	Recommendations
H. N. Gadbail et al. (2026)	AI-enhanced cloud computing	Automation, predictive analytics, chatbots, AIOps	Improved productivity, reliability, and decision-making in cloud environments	Broad overview with limited focus on autonomous cloud operations	Explore deeper integration of Generative AI and autonomous AIOps for cloud-native infrastructures
V. Raj (2025)	AIOps for hybrid cloud operations	Predictive analytics, AIOps	Improved proactive operations and reduced unexpected failures	Limited autonomous decision-making capabilities	Integrate autonomous remediation and self-healing mechanisms
D. K. Seth, K. K. Ratra, and A. P. Sundareswaran (2025)	AI and GenAI for cloud optimization	AI-driven scaling, GenAI automation	Enhanced resource utilization and operational efficiency	Limited discussion of explainability and governance	Develop trustworthy and explainable AI-driven cloud operations
W. C. Potts and C. Carver (2024)	Enterprise and government	Data analytics, IT operations	Highlighted organizational	Focused more on adoption	Investigate scalable autonomous AIOps

	AIOps adoption	integration	factors affecting AIOps success	challenges than automation	frameworks
P. K. Thota (2024)	Autonomous cloud infrastructure management	GenAI, Reinforcement Learning, IaC	Improved scalability and resource management	Limited validation in large-scale real-world environments	Evaluate framework performance in complex cloud-native ecosystems
S. Sundar Ray(2023)	Autonomous incident response	RAG, vector embeddings, agentic workflows	Reduced MTTR and manual intervention	Limited focus on multi-cloud and heterogeneous environments	Extend agentic AI solutions for multi-cloud operations
A. Hayes et al. (2022)	Automated log analysis and root-cause triage	LLMs, RAG, evidence mapping	Faster diagnostics and incident resolution	Limited integration with autonomous remediation systems	Combine GenAI analytics with closed-loop operational automation

6. Conclusion and Future Work

The rapid evolution of cloud-native solutions has transformed how IT operations are managed and operated in today's environment, leading to the need for intelligent, automated, and resilient frameworks. The review focuses on the opportunity of Autonomous AIOps and Generative AI to address the challenges in cloud-based environments. The findings underscore the transformative power of AI-driven systems monitoring and anomaly detection, predictive maintenance, root cause analysis, and intelligent incident management in improving system reliability, operational efficiency, resource usage, and service availability. Furthermore, Generative AI enables automation of log analysis, decision support, and workflows, streamlining operational intelligence and reducing manual effort and accelerating issue resolution. The reviewed studies show that there is a definite move from reactive to proactive and autonomous cloud management. But challenges like explainability, security, privacy, model drift, and complexity of integration will continue to influence the widespread implementation of these enhancements. In conclusion, Autonomous AIOps and Generative AI make for a solid bedrock of intelligent cloud operations for the future.

Future research should aim to build explainable, secure, trustworthy autonomous cloud operation frameworks based on advanced Generative AI/AIOps. Self-healing systems, multi-cloud management, AI governance, and scalable automation are just a few of the areas that still have a lot to be explored in order to build fully autonomous cloud-native environments.

References

[1] R. Lingam, S. Nagi, M. Chigurupati, and B. T. Myneni, “Resilient DevSecOps: Self-Healing Cloud-Native Systems via SRE-Driven AI Threat Detection and Response,” in *2026 International Conference on Smart Futuristic Technology*, IEEE, Jan. 2026, pp. 1–6. doi: 10.1109/ICSFT66733.2026.11508049.

[2] S. Deng *et al.*, “Cloud-Native Computing: A Survey From the Perspective of Services,” *Proc. IEEE*, vol. 112, no. 1, pp. 12–46, 2024, doi: 10.1109/JPROC.2024.3353855.

[3] M. R. C. Mukkolakkal, “InfraLLM: A Generic Large Language Model Framework for Production-Grade Microservice Auto-Scaling in Cloud Infrastructure,” *Int. J. Sci. Res. Mod. Technol.*, vol. 4, no. 11, pp. 113–123, 2025, doi: 10.38124/ijsrmt.v4i11.1023.

[4] H. Allam, “From Monitoring to Understanding: AIOps for Dynamic Infrastructure,” *Int. J. AI, BigData, Comput. Manag. Stud.*, vol. 4, no. 2, pp. 77–86, 2023, doi: 10.63282/3050-9416.IJAIBDCMS-V4I2P109.

[5] M. G. Paolo Notaro, Jorge Cardoso, “A Systematic Mapping Study in AIOps,” 2020. doi: https://doi.org/10.48550/arXiv.2012.09108.

[6] V. S. R. Narapareddy, “Generative AI and Foundation Models,” *Univers. Libr. Innov. Res. Stud.*, vol. 02, no. 02, pp. 07–21, May 2025, doi: 10.70315/uloap.ulirs.2025.0202002.

[7] B. P. Singh and H. Singh, “Using LLMs for Autonomous Cloud Infrastructure Entitlement Management to Prevent Overprivileged Access,” *J. Eng. Comput. Sci.*, vol. 5, no. 4, pp. 1–14, April, 2026, doi: https://doi.org/10.5281/zenodo.19488212.

[8] Y. Patel, “Self-Adaptive Al-Based Orchestration for Multi-Cloud Interoperability and Performance Optimization,” in *SoutheastCon 2026*, Huntsville, AL, USA: IEEE, 2026, pp. 01–08, April. doi: 10.1109/SoutheastCon63549.2026.11476031.

[9] A. K. Padhy, T. P. Patel, V. Soni, A. K. Elengovan, G. B. Thokala, and N. Seshagiri, “Cloud-Native Multimodal Semantic Search and Recommendation for Large-Scale Digital Commerce,” in *2026 4th Odisha International Conference on Electrical Power Engineering, Communication and Computing Technology (ODICON)*, 2026, pp. 1–6, February. doi: 10.1109/ODICON66687.2026.11470613.

[10] Oyekunle Claudius Oyeniran, Adebunmi Okechukwu Adewusi, Adams Gbolahan Adeleke, Lucy Anthony Akwawa, and Chidimma Francisca Azubuko, “Microservices architecture in cloud-native applications: Design patterns and scalability,” *Comput. Sci. IT Res. J.*, vol. 2, no. 1, pp. 92–106, 2024, doi: 10.51594/csitrj.v5i9.1554.

[11] A. Balalaie, A. Heydarnoori, and P. Jamshidi, "Microservices Architecture Enables DevOps: an Experience Report on Migration to a Cloud-Native Architecture," IEEE, 2016, pp. 42–52, March. doi: 10.1109/MS.2016.64.
[12] [12]P. Naayini and S. Kamatala, "Automating Infrastructure Platforms with Cloud, Kubernetes, and Site Reliability Engineering," *Int. J. Comput. Tech.*, vol. 8, no. 6, pp. 1–9, 2021.
[13] V. U. Ugwueze, "Cloud Native Application Development: Best Practices and Challenges," *Int. J. Res. Publ. Rev.*, vol. 5, no. 12, pp. 2399–2412, Dec. 2024, doi: 10.55248/gengpi.5.1224.3533.
[14] A. K. Padhy, C. Medicherla, B. Vulugundam, C. Kulkarni, T. P. Patel, and S. Shivam, "Latency-Optimised Microservices Orchestration for Real-Time E-Commerce in Multi-Cloud Environments," in *2025 International Conference on Computer and Applications (ICCA)*, IEEE, Dec. 2025, pp. 1–6. doi: 10.1109/ICCA66035.2025.11430930.
[15] S. P. P. -, "Cloud Intelligence and AIOps Integration: A Framework for Autonomous IT Operations in Modern Cloud Environments," *Int. J. Multidiscip. Res.*, vol. 6, no. 6, pp. 1–10, 2024, doi: 10.36948/ijfmr.2024.v06i06.33643.
[16] P. Kumar and H. Shreyas Kumar, "AIOps: Analysing Cloud Failure Detection Approaches for Enhanced Operational Efficency," in *2023 International Conference on Applied Intelligence and Sustainable Computing (ICAISC)*, IEEE, Jun. 2023, pp. 1–6. doi: 10.1109/ICAISC58445.2023.10199929.
[17] J. Zhaoxue, L. Tong, Z. Zhenguo, G. Jingguo, Y. Junling, and L. Liangxiong, "A Survey On Log Research Of AIOps: Methods and Trends," *Mob. Networks Appl.*, vol. 26, pp. 2353–2364, 2021, doi: 10.1007/s11036-021-01832-3.
[18] C. Patel, "A Review of Multi-Channel CRM Strategies Using Big Data and Cloud Integration," *Int. J. Sci. Res. Comput. Sci. Eng. Inf. Technol.*, vol. 8, no. 1, January-February, pp. 577–588, 2022, doi: https://doi.org/10.32628/IJSRCSEIT.
[19] A. B. Nassif, M. A. Talib, Q. Nasir, and F. M. Dakalbab, "Machine Learning for Anomaly Detection: A Systematic Review," *IEEE Access*, vol. 9, pp. 78658–78700, 2021, doi: 10.1109/ACCESS.2021.3083060.
[20] Z. Zhong *et al.*, "A Survey of Time Series Anomaly Detection Methods in the AIOps Domain," Aug. 2023. doi: https://doi.org/10.48550/arXiv.2308.00393.
[21] D. K. D. Pal, V. R. R. Alluri, S. Thota, V. S. M. B. S. Chitta, M. Shaik, and C. Schwab, "AIOps : Integrating AI and Machine Learning into IT Operations," *Aust. J. Mach. Learn. Res. Appl.*, vol. 4, no. 1, pp. 288–311, 2024.
[22] T. P. Patel, "Adaptive Token Routing for Heterogeneous LLM Inference in Edge-Cloud Continuum," in *SoutheastCon 2026*, Huntsville, AL, USA: IEEE, 2026, pp. 1–7, April. doi: 10.1109/SoutheastCon63549.2026.11476596.
[23] R. Vadisetty *et al.*, "Generative AI for Cloud Infrastructure Automation," *Int. J. Artif. Intell. Data Sci. Mach. Learn.*, vol. 1, pp. 15–20, 2020, doi: 10.63282/3050-9262.IJAIDSML-V1I3P103.
[24] H. N. Gadbail, K. Sharma, R. Rewatkar, D. Murle, S. Zade, and P. Waghmare, "Integrating AI with Cloud Computing: Architecture, Benefits, and Emerging Innovations," in *2026 7th International Conference on Mobile Computing and Sustainable Informatics (ICMCSI)*, Goathgaun, Nepal, 2026, pp. 1242–1248, March. doi: 10.1109/ICMCSI67283.2026.11412712.
[25] V. Raj, "Utilizing AIOps for Predictive Maintenance in Hybrid Cloud Environments," in *2025 IEEE International Conference on Joint Cloud Computing (JCC)*, 2025, pp. 131–136. doi: 10.1109/JCC67032.2025.00022.
[26] D. K. Seth, K. K. Ratra, and A. P. Sundareswaran, "AI and Generative AI-Driven Automation for Multi-Cloud and Hybrid Cloud Architectures: Enhancing Security, Performance, and Operational Efficiency," in *2025 IEEE 15th Annual Computing and Communication Workshop and Conference (CCWC)*, 2025, pp. 784–793. doi: 10.1109/CCWC62904.2025.10903928.
[27] W. C. Potts and C. Carver, "Best Practices Implementing AIOps in Large Organizations," in *2024 International Conference on Smart Applications, Communications and Networking (SmartNets)*, IEEE, May 2024, pp. 1–5. doi: 10.1109/SmartNets61466.2024.10577643.
[28] P. K. Thota, "A Generative AI Framework for Autonomous Infrastructure Management in Cloud Operations," *Int. Sci. J. Eng. Manag.*, vol. 03, no. 10, pp. 1–8, Oct. 2024, doi: 10.55041/ISJEM02027.
[29] S. Sundar Ray, "Autonomous Incident Response Using Generative AI and Agentic Systems in Distributed Enterprise Architectures," *Available SSRN 6647338*, 2023.
[30] A. Hayes, E. Carter, D. Foster, S. Reynolds, M. Bennett, and J. Krishnan, "From Logs to Insights: Generative AI for Automated Root-Cause Triage in Distributed Enterprise Systems," 2022. doi: 10.5281/zenodo.20265420.