EVALUATING AI-DRIVEN AUTOMATION TECHNIQUES IN CLOUD INFRASTRUCTURE MANAGEMENT

by

MARCEL AGYEBIT KOHBAH
B.S., University of Bamenda, 2016
M.S., University of Maryland Global Campus, 2020

A Research Paper Submitted to the School of Computing Faculty of
Middle Georgia State University in
Partial Fulfillment of the Requirements for the Degree

DOCTOR OF SCIENCE IN INFORMATION TECHNOLOGY

MACON, GEORGIA
2025

Evaluating AI-Driven Automation Techniques in Cloud Infrastructure Management

Marcel Kohbah, Student, marcel.kohbah@mga.edu

Abstract

This study examines how AI-driven automation improves resource provisioning, cost optimization, and overall performance in cloud infrastructure management. Through a qualitative analysis of technical reports and peer-reviewed articles from 2014 to 2024, it explores how machine learning, deep reinforcement learning, and predictive analytics enhance efficiency and reduce downtime. This research adopts a qualitative approach, utilizing thematic analysis to systematically uncover, examine, and interpret recurring patterns (themes) in existing literature on the role of AI-driven automation in managing cloud infrastructure. Findings indicate that AI-driven solutions streamline operations and bolster security through real-time threat detection, yet challenges persist in data privacy, integration complexity, and model transparency. Based on these insights, the research offers practical recommendations for organizations adopting AI-driven automation, providing a roadmap for maximizing scalability and innovation in modern cloud environments.

Keywords: AI-driven automation, cloud infrastructure management, anomaly detection, machine learning, cost optimization, DevOps, IT operations

Introduction

Cloud computing and Al are distinct and complementary technologies that offer tremendous economic and consumer benefits. The cloud reduces cost and democratizes access to computational resources, allowing Al to streamline business functions and provide new insights that improve consumer welfare (United States Congress House, 2020). Cloud computing equips artificial intelligence (AI) with tremendous power and is considered one of the most critical catalysts for developing innovative, intelligent applications. With its potential to change how data is stored and processed across various geographies, AI's scope and impact have reached a larger market. With all the cloud models, AI developers and consumers have started to create an ecosystem that improves the lives of millions (Mohammed et al., 2023). This research seeks to bridge this knowledge gap by systematically evaluating the effectiveness of AI-driven automation techniques in cloud infrastructure management. Through a comprehensive analysis, this study will provide valuable insights into how AI technologies can be harnessed to improve cloud operations, ultimately offering evidence-based recommendations for organizations contemplating the adoption of AI-driven automation solutions. This research aims to contribute to the broader discourse on the role of AI in shaping the future of cloud computing and IT infrastructure management.

Problem statement

Cloud computing, a new form of time-sharing that emerged just a decade ago, has brought about a paradigm shift in the IT landscape. This technology allows a network of computer resources to be linked together through the internet and deployed to meet users' needs as they occur, expanding the computing power exponentially. However, this new frontier of cloud systems management has posed significant challenges for IT professionals, necessitating effective solutions (Kenninger, 2019). AI and cloud computing converge in automating data analysis, data management, security, and decision-making processes. The ability of AI to exercise machine learning and to derive impartial interpretations of data-driven insights fuels efficiency in these processes and can lead to significant cost savings on numerous fronts within the enterprise (Brenner, 2023). Cloud computing is not just an enabler but a foundation for AI development, offering the flexibility and accessibility essential for the efficient development and implementation of AI applications.

The interaction between these technologies allows the training of AI models on a large scale, accelerating solutions' development and improvement cycle. Additionally, cloud computing democratizes access to new AI applications, eliminating barriers for companies of all sizes and driving innovation at an unprecedented rate (Tchule, 2024). Currently, researchers are facing a huge technological shift in the direction of core computing domains. The reasons for merging technology are to overcome the limitation of the solution to solve the problems on a large scale. In the artificial intelligence era, the IoT is making more contributions by making different utilities equipped with AI and ML concepts (Pandey et al., 2023). Leveraging AI automation with cloud computing resource allocation is needed to enhance cloud infrastructure management.

Purpose of the study

The purpose of this study is to assess the effectiveness of AI-driven automation techniques in cloud infrastructure management. Its objectives are to provide a comprehensive understanding of the benefits and challenges associated with AI-driven automation in cloud infrastructure management, thereby guiding organizations in making informed decisions about its adoption and implementation with cloud-native applications. This study aims to contribute to the broader discussions on AI and cloud computing by elucidating how integrating AI technologies enhances operational efficiency and scalability and drives innovation in cloud environments.

Research question

RQ1: Based on the reviewed literature, what key techniques can be identified in AI-driven automation for cloud infrastructure management to determine how these techniques can optimize resource provisioning efficiency?

Research objectives

The research will evaluate the key techniques of AI-driven automation in cloud infrastructure management and assess their effectiveness in optimizing resource provisioning efficiency. By investigating how AI technologies enhance cloud management, this research will provide evidence-based recommendations for organizations considering the adoption of AI-driven automation. Additionally, the study seeks to contribute to the broader discourse on the convergence of AI and cloud computing in shaping the future of IT infrastructure management and innovation.

Proposal organization

This research proposal is organized as follows. A preliminary literature review is presented, followed by a proposed methodology that includes a procedure description and data analysis.

Review of the Literature

Integrating AI in cloud computing transforms various industries by enhancing efficiency, optimizing resources, and improving performance. Studies highlight AI's role in cloud infrastructure, automotive industry advancements, cybersecurity enhancements, and Site Reliability Engineering. AI-driven automation and predictive analytics in cloud environments enable better resource allocation, proactive maintenance, and faster incident responses, leading to significant cost savings and operational efficiencies. This convergence promises a future of sustainable, efficient, and innovative solutions across multiple sectors.

Efficiency Gains and Operational Optimization

AI-driven automation transforms cloud computing, electric vehicles (EVs), and IT infrastructure management by enhancing efficiency, resource optimization, and system resilience. Traditional manual methods need to be improved for the dynamic nature of cloud environments. AI automation, employing techniques like machine learning, evolutionary algorithms, and deep reinforcement learning, offers promising solutions. AI-driven methods can significantly improve resource utilization, energy efficiency, and task scheduling. For example, deep reinforcement learning-based approaches, such as actor-critic methods, have outperformed traditional methods in resource allocation, reducing operational costs and enhancing performance (Karamthulla et al., 2023).

This AI-driven efficiency extends beyond cloud environments into transportation, where EVs benefit from cloud-based data analytics and AI-powered automation. EVs benefit from AI and cloud-based data analytics, enhancing performance, energy efficiency, and user experience. Critical AI applications include dynamic charging optimization, predictive maintenance, and intelligent fleet management. Battery technology, vehicle design, and charging infrastructure have significantly advanced. Cloud computing supports remote diagnostics and over-the-air updates, while AI enhances autonomous driving and energy management. This convergence promises a future of sustainable mobility, efficient energy use, and innovative transportation solutions (Brian, 2023).

Similarly, AI Operations (AIOps) revolutionize cloud infrastructure by enhancing predictive analytics, automating incident response, and ensuring system reliability. Revolutionizing cloud infrastructure management through AIOps and significantly enhancing predictive analytics, automated incident response, and system reliability. AI algorithms preemptively identify potential disruptions, optimizing resource use and reducing downtime, while automated incident responses accelerate problem resolution without human intervention. These AI-driven processes ensure high availability and system resilience, crucial for continuous service delivery. The deployment of AIOps improves operational efficiency and reliability and scales seamlessly with growing infrastructures, providing substantial cost savings and enhancing customer satisfaction through stable, efficient service delivery (Infuwiki, 2024).

Security Enhancements and Threat Mitigation

AI-driven automation has transformed security operations centers (SOCs) by enhancing efficiency and effectiveness in cybersecurity. SOCs face data overload, talent shortages, and increasing attack sophistication. AI offers solutions by automating routine tasks, reducing false positives, and enabling faster, more accurate threat detection and response. Critical AI applications include threat detection using machine learning, alert prioritization, and incident response through AI-driven playbooks and SOAR technologies. Successful AI adoption in SOCs requires workflow integration, transparent AI, and continuous model validation to augment human analysts, improving overall security operations and resilience against cyber threats (Zhang, 2024).

Beyond SOCs, AI-driven automation optimizes cloud infrastructure by handling resource provisioning, troubleshooting, and predictive maintenance. AI-driven automation handles resource provisioning and troubleshooting, reducing manual intervention. Predictive analytics foresees potential issues, enabling proactive maintenance. AI optimizes CI/CD processes, enhances security with real-time threat detection, and improves cost management by analyzing usage patterns. Additionally, AI-based performance monitoring identifies bottlenecks and suggests optimizations, while NLP enhances communication within DevOps teams. This synergy between AI and cloud computing significantly improves cloud-based systems' efficiency, scalability, and reliability (Akhilapemmaraju, 2024).

AI also enhances cloud applications through intelligent resource allocation, predictive scaling, and anomaly detection. In his article, Kunduru examines how AI enhances cloud application efficiency. It explores AI-driven techniques, such as resource allocation, predictive scaling, and anomaly detection, which help improve operational responsiveness and performance reliability. The author highlights how machine learning and deep learning models analyze historical data to anticipate demand, enabling real-time adjustments that prevent over or under-provisioning. While AI presents notable benefits, challenges like data bias and privacy concerns persist. The study underscores AI's transformative potential in cloud environments, promising further optimization advancements as AI technologies mature (Kunduru, 2023).

Moreover, AI-driven cloud solutions streamline business processes, enhance customer insights, and improve decision-making. Lakshmisri Surya's examines the role of artificial intelligence (AI) in enhancing cloud applications, underscoring its benefits for organizational efficiency and cost-saving. By integrating AI with cloud technologies, companies gain streamlined order-to-cash processes, enhanced customer insights, and improved decision-making. AI's predictive capabilities allow cloud systems to analyze vast datasets, providing insights into customer behaviors and operational patterns. Moreover, AI-driven automation supports back-office functions, reducing errors and increasing productivity. The article highlights the potential of AI to bolster security, mitigate unauthorized access, and enhance data recovery. The article illustrates AI's transformative impact on cloud environments, proposing that it is integral for future business growth and innovation (Surya, 2018).

Implementation Challenges and Research Gaps

The convergence of Artificial Intelligence (AI) and cloud computing is revolutionizing traditional industries by fostering innovative and transformative changes. This integration enables harnessing AI's capabilities, such as learning, reasoning, and problem-solving, alongside the scalability and remote accessibility provided by cloud computing. Together, they facilitate a dynamic environment where computational power and data handling are significantly enhanced, paving the way for advanced applications in various fields. The symbiotic relationship enhances operational efficiencies, drives cost savings, and democratizes access to cutting-edge technology, catalyzing innovation across sectors (Team_Qwikskills, 2023).

Despite its potential, the AI-cloud ecosystem presents infrastructure flexibility, cost management, security, and performance monitoring challenges. Mohammed, Fang, and Ramos discuss how AI-powered cloud models, including Infrastructure as a Service (IaaS), Platform as a Service (PaaS), and Software as a Service (SaaS), enhance cloud efficiency. Critical applications discussed include autonomous operations, predictive analytics, CI/CD optimization, cost management, security enhancement, and performance monitoring. The issue also presents innovative research on power measurement frameworks, educational performance analytics, object coreference resolution, multimedia similarity ranking, CPU load prediction, and intelligent healthcare systems, emphasizing AI's role in driving strategic and operational advancements in cloud environments (Mohammed et al., 2021).

One critical challenge in AI-cloud integration is ensuring service reliability. Wen and Zheng examine the integration of artificial intelligence (AI) and cloud computing, focusing on AI's impact on enhancing cloud services and addressing reliability issues. Highlighting AI's transformative role in various fields, the authors propose a reliability verification method at the IaaS layer, using performance metrics such as QPS, CPU, and memory. Experimental results demonstrate the method's effectiveness, offering a structured approach for assessing cloud service reliability, which is vital for users selecting cloud providers. This research underscores the importance of AI in optimizing cloud infrastructure and offers a practical solution for reliability assessment in cloud environments (Wen & Zheng, 2021).

Additionally, integrating AI with distributed systems presents opportunities and challenges for cloud computing. In their article, Zangana and Zeebaree review AI-driven applications and their impact on

performance optimization, security, and IoT integration. It reviews recent developments in AI-driven applications, examining contributions across performance optimization, security, and IoT integration. The authors provide insights into how AI and distributed systems synergize to address cloud computing challenges by evaluating foundational theories and practical applications. The article further offers comparative analyses of different approaches. It suggests future research directions to advance distributed AI systems, ultimately acting as a resource for professionals and researchers in cloud-based AI innovations (Zangana & Zeebaree, 2024).

Methodology

Textual analysis has long been a staple of academic, scientific, and industrial research. This process of dissecting, interpreting, and deriving insights from text data is instrumental in understanding complex phenomena within various domains such as social sciences, humanities, computer science, and linguistics (Milev, 2023). This study employs a qualitative research design using thematic analysis to identify, analyze, and interpret patterns (themes) within published literature on AI-driven automation in cloud infrastructure management. Thematic analysis involves assigning data to several codes, grouping codes into themes, and then identifying patterns and interconnections between these themes. A thematic analysis allows a nuanced understanding of what people say and do within their social contexts (Jowsey et al., 2021). Thematic analysis was chosen for its flexibility in systematically identifying and interpreting patterns across diverse texts, enabling a comprehensive understanding of complex phenomena. Its capacity to organize data into meaningful themes provides nuanced insights into the contextual factors influencing AI-driven automation techniques in cloud infrastructure management. The qualitative nature of this research will focus on interpreting meanings, examining contextual factors, and understanding the implications of AI-driven strategies on cloud infrastructure management.

Data Collection

The selected texts were systematically gathered through academic databases (e.g., IEEE Xplore, Google Scholar, ACM Digital Library, MGA Library, and public libraries), industry sources (e.g., Gartner, Forrester), and cloud service provider publications from popular cloud providers like AWS, AZURE, and GCP. The below library search query was used to search for already existing articles and journals from various libraries:

- ("AI-driven automation" OR "AI-based automation" OR "intelligent automation" OR "machine learning") AND ("cloud infrastructure management" OR "cloud management" OR "cloud computing" OR "cloud service management" OR "cloud orchestration") AND ("effectiveness" OR "evaluation" OR "impact" OR "performance") AND ("automation tools" OR "risk mitigation" OR "compliance strategies" OR "predictive analytics" OR "cloud cost optimization")

Data analysis

The data for this research will consist of technical reports, industry white papers, and peer-reviewed articles related to AI-driven automation in cloud infrastructure management, published between 2014 and 2024. The research will focus on texts that discuss the implementation, benefits, challenges, and future of AI-driven automation in cloud management. The analysis is structured around several core themes that emerged from the data, including DevOps, Operational Efficiency, Anomaly Detection, Security, Machine Learning, and Cost Optimization. Each theme is discussed in detail with related sub-themes, highlighting the contributions of AI-driven automation techniques to cloud infrastructure management.

The analysis process uses a thematic analysis approach. Virginia Braun and Victoria Clarke originally developed this process for psychology research. However, thematic analysis is a flexible method that can be adapted to many kinds of research (Braun & Clarke, 2006).

Starting with the data familiarization step, relevant documents were reviewed to understand how they address AI-driven automation in cloud management. A research log noted key points, ideas, and recurring concepts (e.g., resource provisioning, cost optimization, security enhancements). Continuing with the next step, the coding process involved systematically reviewing each selected text for significant statements. Codes were assigned to segments of text describing AI-driven techniques, outcomes, barriers, or future directions. These codes were grouped based on conceptual similarities, highlighting broader categories such as resource allocation, predictive analytics, system reliability, and integration challenges. In coding, short phrases or words, known as codes, are assigned to data segments that capture the core message, significance, or theme of the data. This step simplifies complex textual data by transforming it into a theoretical form and assists in identifying elements related to the research questions. Keywords are essential in coding as they form the backbone of the analysis and help convert raw data into insightful, manageable units (Naeem et al., 2023).

On to the next step, generating themes, categories from the axial coding stage were examined to uncover overarching themes (e.g., "Efficiency Gains Through AI Automation," "Security and Compliance," "Integration and Scalability Challenges"). Each theme's relevance and distinctness were checked against the entire dataset. Overlapping themes were merged or subdivided for clarity. Themes were clearly defined and labeled to capture their central organizing concept. When reviewing themes, each theme was interpreted in context, relating key findings to the research questions:

- Based on the reviewed literature, what key techniques can be identified in AI-driven automation for cloud infrastructure management to determine how these techniques can optimize resource provisioning efficiency?

A rigorous process of checking the integrity of these themes involved reading and re-reading their data and checking to see if the themes 'fit' the coded excerpts. Data sets were compared against different select themes. Themes were clearly defined and labeled to capture their central organizing concept. After checking the themes and including any additional emerging themes, final names were set for themes identified. Each theme was listed after formulating precisely what each theme was and figuring out how it helped understand the data. Once coding and theme development were completed, each theme was presented clearly and logically. It began by outlining the central idea behind each theme, followed by its subthemes or relevant dimensions. A conclusion explained the main takeaways and how the analysis answered our research question (Braun & Clarke, 2006; Naeem et al., 2023).

Results

The results of this study highlight the transformative impact of AI-driven automation techniques on cloud infrastructure management, particularly in areas such as DevOps, operational efficiency, anomaly detection, security, machine learning, and cost optimization. By integrating AI technologies into cloud environments, organizations achieve significant reliability, efficiency, and security improvements while addressing resource allocation, scalability, and cost management challenges. Figure 1 below shows a thematic map with themes, sub-themes, links, and their relations.

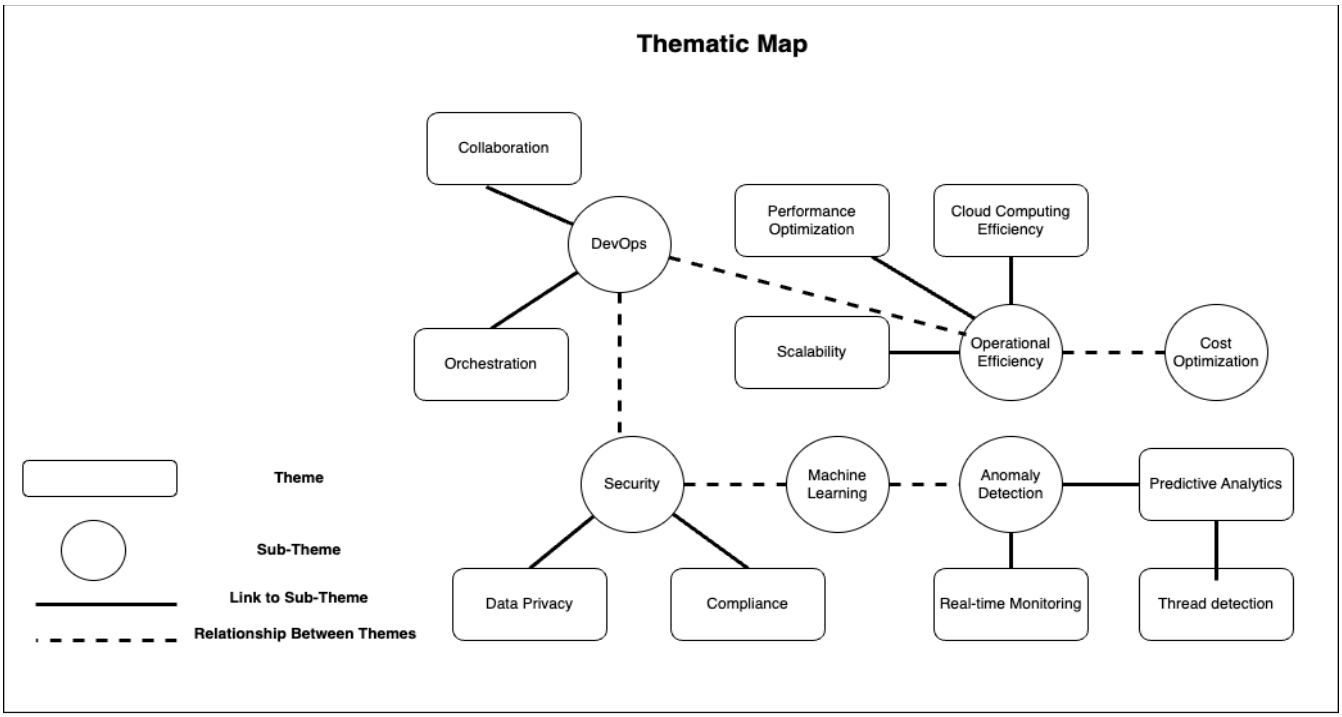

Figure 1: Thematic Map

DevOps

AI technologies are increasingly integrated into DevOps practices to enhance cloud application reliability, streamline workflows, and improve collaboration. AI technology builds cloud applications' reliability by providing solutions that ensure organization continuity, efficient data backup, and more comfortable recovery from disaster. Moreover, AI technology improves existing cloud application solutions and develops new paths that are appropriate for development. AI-driven playbooks automate and orchestrate investigation workflows across disparate systems and data sources, enabling SOC teams to uncover stealthy attacks and automate remediation actions where applicable. AI project themes can use AI collaboration tools in their application development lifecycle.

Operational Efficiency

AI-driven automation techniques significantly enhance operational efficiency by minimizing resource wastage, reducing operational costs, and promoting environmental sustainability through lower energy consumption. AI algorithms help automate complex security operation center workflows. Moreover, AI tools enhance the real-time detection of security vulnerabilities by detecting areas needed for improvement. AI can predict resource allocation of cloud applications based on usage by scaling resources during high and low usage, thereby increasing performance optimization**.** When leveraged with cloud computing, AI can investigate data patterns and optimize cloud resource workflow by enhancing efficiency**.** AI solutions let IT operation teams scale compute resources depending on peak and low usage demands without human intervention to address resource scalability issues.

Anomaly Detection

AI-based anomaly detection serves as a proactive defense mechanism in cloud environments, continuously monitoring system metrics and swiftly responding to deviations from normal behavior. AI solutions can identify command and control communications, data exfiltration attempts, and malware delivery in progress when integrated with real-time monitoring. It provides predictive analytics of cloud resource usage based on historical data available to AI solutions. By implementing AI-driven automation techniques, security operation centers have realized significant benefits, including higher threat detection rates, faster incident response times, and more efficient analyst workflows.

Security

AI plays a crucial role in cloud application security by analyzing data to identify potential inconsistencies and threats. AI-driven automation solutions transform SOC operations, allowing organizations to protect their assets against evolving cyber threats. AI also assists human security teams by improving their ability to recognize and respond effectively to security challenges. AI PaaS empowers AI automation techniques to protect data and models, enabling non-machine learning users to create models with data privacy. AI systems are increasingly employed to process product and service payments in accordance with industry standards.

Machine Learning

Machine learning and deep learning algorithms are essential to AI-driven resource allocation. They facilitate data-informed decision-making and real-time adjustments based on historical usage patterns and current demand. AI automation can predict the resource utilization of cloud applications by optimizing resource allocations. Algorithms like machine learning and deep learning leverage historical data to forecast usage patterns and allocation. These algorithms improve resource allocation by managing complex data representations and learning intricate patterns, ultimately enhancing cloud application performance.

Cost Optimization

AI-driven cost optimization strategies analyze usage patterns to recommend cost-saving measures such as rightsizing resources, scheduling workloads during off-peak hours, and effectively utilizing reserved instances. These techniques tackle the challenges associated with under-provisioning resources in cloud infrastructure. With most cloud service providers adopting a pay-as-you-go model, AI enables organizations to pay only for their actual resource needs. AI algorithms evaluate the usage patterns of cloud resources and propose cost-reduction strategies to combat both over- and underutilization of these resources.

Discussion of Findings

The findings from this study illustrate the effect of AI-driven automation techniques on cloud infrastructure management, directly addressing the research question (RQ1): Based on the reviewed literature, what key techniques can be identified in AI-driven automation for cloud infrastructure management to optimize resource provisioning efficiency? The analysis reveals that machine learning, deep reinforcement learning, and predictive analytics are crucial for enhancing operational efficiency, security, and cost optimization. For example, the "Machine Learning" and "Operational Efficiency" themes emphasize how AI algorithms predict usage patterns and dynamically scale resources to ensure optimal performance during fluctuating demand. The findings align with the literature (e.g., Karamthulla et al., 2023), indicating that AI-driven methods outperform traditional resource allocation approaches.

The "DevOps" and "Anomaly Detection" themes demonstrate AI's role in streamlining workflows and proactively mitigating risks, such as real-time threat detection and automated incident responses. These capabilities reduce downtime, a persistent challenge in cloud systems, and significantly improve the efficiency, scalability, and reliability of cloud-based systems (Akhilapemmaraju, 2024 and Kenninger, 2019). Similarly, the "Security" theme underscores AI's transformative effect on cybersecurity, with threat detection enhancing response times in security operations centers (Zhang, 2024). The "Cost Optimization" theme further reveals practical strategies, such as rightsizing and workload scheduling, that leverage AI to align costs with actual needs, addressing over- and under-provisioning issues.

Implications of Findings

The implications of these findings indicate that organizations can adopt AI-driven techniques like predictive scaling and anomaly detection to optimize resource provisioning, lower operational costs, and enhance security, as illustrated by the results. This is especially pertinent for organizations utilizing cloud providers such as AWS, Azure, and GCP, which operate on a pay-as-you-go model. The focus on real-time threat detection and automated responses also suggests that AI can bolster cybersecurity resilience, a vital necessity given the increasing sophistication of attacks (Zhang, 2024). This study supports the convergence theory of AI and cloud technologies, suggesting a need to refine models to include human capital variables. Organizations should invest in training their staff on AI, which could potentially boost adoption based on our findings.

Conclusion

This research concludes that AI-driven automation techniques in cloud infrastructure management, including machine learning, predictive analytics, DevOps, and anomaly detection, significantly enhance management by optimizing resource provisioning efficiency, improving security, and reducing costs. The thematic analysis addresses RQ1 by identifying key techniques that enable dynamic resource scaling, proactive anomaly detection, and usage-based cost optimization, as seen across the themes of "Operational Efficiency," "Security," and "Cost Optimization." These advancements align with the study's objective to provide evidence-based insights for organizations, confirming AI's role in driving operational excellence and innovation in cloud environments.

Integrating artificial intelligence with cloud automation has emerged as a transformative force in modern infrastructure management, demonstrating significant impacts across various industrial sectors. The comprehensive analysis of AI-driven predictive scaling mechanisms, autonomous operations, and advanced container orchestration systems has established that artificial intelligence is crucial in enhancing operational efficiency and system reliability (Vanam, 2025).

Despite these advantages, several challenges remain, such as data privacy concerns, integration complexity, and issues with model interpretability. Organizations must navigate these challenges carefully to maximize the benefits of AI automation in cloud management. The study identifies skill gaps in AI adoption as a limiter, offering a foundation for future research. Future studies could explore AI ethics in cloud computing management.

Limitations of the Study

A qualitative approach that relies on thematic analysis of literature from 2014 to 2024 lacks quantitative validation of the effectiveness of AI-driven techniques (e.g., precise metrics on cost savings or downtime reduction). This limitation restricts the ability to measure real-world impact beyond reported trends.

However, qualitative research is often considered to be lacking scientific rigor. Unfortunately, the standard strategies used to enhance validity, reliability, and objectivity in quantitative research are not always relevant to qualitative research (Verhoef & Casebeer, 1997). AI technologies in cloud computing are rapidly evolving, so some findings may become outdated as new AI models and automation techniques emerge. The study used existing literature, primarily qualitative rather than quantitative, which would have provided a real-time example. The research relies heavily on old articles. To address this, future research could incorporate surveys and actual data from cloud companies, utilizing the most recent information from the past two years. It should include more small businesses from various areas and conduct more extended studies to observe how AI usage in cloud computing management changes over time using a quantitative approach.

Recommendations for Future Research

To address these limitations, future research should focus on quantitative studies, such as experiments or case studies, to measure the tangible impacts of AI-driven automation (e.g., percentage improvements in resource efficiency or security incident response times). Incorporating primary data from real-world cloud deployments would enhance the empirical foundation. Future research can also focus on the ethical considerations when using AI in cloud computing resource management by reviewing existing gaps.

References

Akhilapemmaraju. (2024, April 14). Advanced cloud computing and devops with ai involve leveraging artificial intelligence (AI) technologies. *Medium*. https://medium.com/@akhilapemmaraju20/advanced-cloud-computing-and-devops-with-ai-involve-leveraging-artificial-intelligence-ai-e94f719463ad

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. Qualitative Research in Psychology, 3(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Brenner, M. (2023, July 12). As more businesses turn to public cloud to leverage artificial intelligence innovation, learn why data privacy becomes more challenging. The Forecast by *Nutanix*. https://www.nutanix.com/theforecastbynutanix/technology/ai-in-the-cloud

Brian, A. D. A. (2023, January 18). AI-driven horizons: electric vehicles, cloud computing, and the future of automation. https://ijaeti.com/index.php/Journal/article/view/180

Infuwiki. (2024, March 9). Evolving cloud infrastructure management with AI: the role of AIOps in site reliability engineering. *Medium*. https://infuwiki.medium.com/evolving-cloud-infrastructure-management-with-ai-the-role-of-aiops-in-site-reliability-engineering-b1106932cb26

Jowsey, T., Deng, C., & Weller, J. (2021). General-purpose thematic analysis: a useful qualitative method for anesthesia research. *BJA Education*, *21*(12), 472–478. https://doi.org/10.1016/j.bjae.2021.07.006

Karamthulla, M. J., Narkarunai, J., Malaiyappan, A., & Tillu, R. (2023). Optimizing resource allocation in cloud infrastructure through AI automation: a comparative study. *Journal of knowledge learning and science technology*, 2(2), 315–326. https://doi.org/10.60087/jklst.vol2.n2.p326

Kenninger, J. (2019, December). Executive insights on cloud computing, automation, and building a more agile IT team. https://research.ebsco.com

Kunduru, A. R. (2023). Artificial intelligence usage in cloud application performance improvement. *Central Asian journal of mathematical theory and computer sciences.* https://cajmtcs.centralasianstudies.org/index.php/CAJMTCS/article/view/491/552

Milev, P. (2023). The role of data visualization in enhancing textual analysis. *Trakia journal of sciences*, 21, 248–253. https://doi.org/10.15547/tjs.2023.s.01.042

Mohammed, S., Fang, W. C., & Ramos, C. (2021b). Special issue on '"artificial intelligence in cloud computing."' Computing, 105(3), 507–511. https://doi.org/10.1007/s00607-021-00985-z

Naeem, M., Ozuem, W., Howell, K., & Ranfagni, S. (2023). A Step-by-Step process of thematic analysis to develop a conceptual model in qualitative research. *International Journal of Qualitative Methods*, *22*. https://doi.org/10.1177/16094069231205789

Pandey, N. K., Kumar, K., Saini, G., & Mishra, A. K. (2023). Security issues and challenges in cloud of things-based applications for industrial automation. *Annals of Operations Research*, *342*(1), 565 584. https://doi.org/10.1007/s10479-023-05285-7

Surya, L. (2018, October 10). Streamlining cloud application with AI technology. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3785667

Tchule, R. (2024, May). AI and cloud computing: an intrinsic link. *automation.com.* https://www.automation.com/en-us/articles/may-2024/ai-cloud-computing-intrinsic-link

Team_Qwikskills. (2023, September 12). The future is now: integrating AI with cloud computing for innovation. *QwikSkills*. https://qwikskills.com/blog/the-future-is-now-integrating-ai-with-cloud-computing-for-innovation/

United States Congress. (2020). AI and the evolution of cloud computing: evaluating how financial data is stored, protected, and maintained by cloud providers. https://financialservices.house.gov/

Vanam, N. G. (2025). AI-enhanced cloud automation: a framework for next-generation infrastructure management. *International journal of scientific research in computer science engineering and information technology*, *11*(1), 12–19. https://doi.org/10.32628/cseit25111204

Verhoef, M. J., & Casebeer, A. L. (1997). Broadening horizons: Integrating quantitative and qualitative research. *Canadian journal of infectious diseases and medical microbiology*, *8*(2), 65–66. https://doi.org/10.1155/1997/349145

Wen, X., & Zheng, Y. (2021, September 1). The application of artificial intelligence technology in cloud computing environment resources. *River publishers journals & magazine | IEEE Xplore*. https://ieeexplore.ieee.org/abstract/document/10247157

Zangana, H. M., & Zeebaree, S. R. M. (2024). Distributed systems for artificial intelligence in cloud computing: a review of AI-powered applications and services. *International journal of informatics information system and computer engineering (INJIISCOM)*, *5*(1), 11–30. https://doi.org/10.34010/injiiscom.v5i1.11883

Zhang, W. C. a. J. (2024, February 6). Elevating security operations: the role of AI-driven automation in enhancing SOC efficiency and efficacy. https://journals.sagescience.org/index.php/jamm/article/view/128