International Journal of Computer Applications Technology and Research
Volume 14–Issue 07, 01 – 11, 2025, ISSN:-2319–8656
DOI:10.7753/IJCATR1407.1001

LLMOps, AgentOps, and MLOps for Generative AI: A Comprehensive Review

Satyadhar Joshi
Independent, Alumnus, International MBA, Bar-Ilan University, Israel

Abstract: AgentOps is an emerging discipline focused on the operationalization, monitoring, and optimization of agentic AI systems, especially in the context of generative AI. This paper provides an overview of AgentOps, its relationship to MLOps and GenAIOps, and best practices for deploying and managing agentic AI in enterprise environments. This paper shows how the evolution happened form operational methodologies in AI, from traditional Machine Learning Operations (MLOps) to finally specialized Large Language Model Operations (LLMOps) and then to Generative AI Operations (GenAIOps). We examine the key challenges in operationalizing generative AI, including model monitoring, prompt management, agent debugging, and ethical considerations. Through a systematic review of over 100 recent articles and industry practices, we identify critical gaps in current operational approaches and summarize based on recent literature an integrated framework that combines MLOps, LLMOps, and AgentOps principles. We present case studies demonstrating successful implementations and provide recommendations for organizations transitioning to generative AI at scale. We delineate the distinct characteristics, challenges, and best practices associated with each stage, emphasizing how AgentOps extends prior concepts to manage the unique complexities of multi-agent systems, including monitoring, debugging, and secure lifecycle management. We provide a holistic overview of the current landscape and future directions for operationalizing advanced AI systems.
Keywords: Generative AI, MLOps, LLMOps, AgentOps, AI Operations, Model Deployment, AI Agents, Agentic AI, GenAI, MLOps, Observability, AI Operations

1. INTRODUCTION

Traditional MLOps and GenAIOps frameworks are evolving to address the unique requirements of agentic AI, giving rise to the field of AgentOps [1], [2], [3].

The emergence of generative AI has fundamentally transformed the artificial intelligence landscape, introducing new operational challenges that extend beyond traditional machine learning workflows [4]. While MLOps established best practices for deploying and monitoring predictive models, generative AI systems require additional considerations for prompt engineering, hallucination mitigation, and multi-agent coordination [5].

This paper makes three primary contributions:

- A systematic analysis of operational requirements for generative AI systems across three dimensions: model operations (LLMOps), agent operations (AgentOps), and traditional machine learning operations (MLOps)
- Summary of current framework that combines these operational paradigms while addressing their unique challenges based on litreature
- Recent updates on guidelines for implementing various framework based on studies and industry best practices and application areas

Generative Artificial Intelligence (GenAI) has marked a transformative shift across industries, especially finance enabling capabilities in content creation, design, scientific research, risk analytics, and decision-making [6], [7]. From image generation to human-like text production and code synthesis and also artificial data, GenAI models are rapidly moving from research labs to production environments. However, their successful deployment and sustained operation present unique and complex challenges that extend beyond traditional Machine Learning Operations (MLOps) in areas like hallucination which are not properly understood. As these models evolve into more autonomous and interactive AI agents, the need for specialized operational frameworks becomes even more critical.

This paper provides a comprehensive overview of the evolving landscape of AI operations, highlighting the progression from foundational MLOps principles to the specialized domains of Large Language Model Operations (LLMOps) and Generative AI Operations (GenAIOps). We then introduce Agent Operations (AgentOps) as the latest frontier, specifically designed to address the distinct requirements for managing, monitoring, debugging, and securing complex AI agent systems. Drawing upon a broad range of contemporary literature and industry insights, we explore the challenges inherent in operationalizing GenAI and agents, and present best practices and solutions for building scalable, reliable, and ethically sound AI deployments.

2. BACKGROUND AND RELATED WORK

AgentOps builds on concepts from MLOps, LLMOps, and GenAIOps, providing tools for observability, debugging, and lifecycle management of AI agents [8], [9]. Platforms such as AgentOps.ai and NexaStack offer dashboards, session replays, and cost tracking for agentic workflows.

2.1 Evolution from MLOps to GenAIOps

Traditional MLOps focuses on the continuous integration and delivery (CI/CD) of machine learning models [10]. However, as noted in [11], generative AI introduces new requirements that necessitate the evolution to GenAIOps (Generative AI Operations). Key differences include:

- Dynamic input/output relationships in generative models [12]
- Need for prompt versioning and management [13]
- Complex evaluation metrics beyond traditional accuracy measures [14]

2.2 LLMOps: Specialized Operations for Large Language Models

LLMOps addresses the unique challenges of deploying and maintaining large language models [15]. As discussed in [16], critical components include:

"LLMOps"="Model Management"+"Prompt Engineering" newline +"Hallucination Monitoring"

2.3 AgentOps: Managing AI Agent Lifecycles

The rise of autonomous AI agents has led to the emergence of AgentOps [17]. As defined in [18], AgentOps encompasses:

- Agent monitoring and debugging [19]
- Multi-agent coordination [20]
- Performance optimization [21]

A typical AgentOps stack includes:

- **Monitoring**: Real-time metrics, logs, and session replays.
- **Debugging**: Time-travel debugging, error tracing, and prompt injection detection.
- **Automation**: Deployment pipelines, scaling, and rollback mechanisms.
- **Security**: Audit trails and compliance reporting.

2.4 Best Practices

- Integrate observability early in the agent development lifecycle.
- Use automated testing and CI/CD for agent updates.
- Monitor agent costs and performance continuously.
- Ensure robust error handling and rollback strategies.

3. BACKGROUND: EVOLUTION OF AI OPERATIONS

3.1 Machine Learning Operations (MLOps)

MLOps serves as the central discipline for operationalizing machine learning models and has been implemented successfully in last 10 years. It integrates DevOps principles—Continuous Integration (CI), Continuous Delivery (CD), and Continuous Training (CT)—into the machine learning lifecycle [10]. The core objective of MLOps is to streamline the process of taking ML models from experimentation to production and maintaining them effectively [22], [23]. Key aspects of MLOps include:

- Data preparation and versioning [24].
- Model training, versioning, and experiment tracking [25].
- Model deployment and inference serving.
- Model monitoring for performance degradation (e.g., data drift, model drift) and retraining automation [26].
- Collaboration across data scientists, ML engineers, and operations teams [27].

Various platforms and services have emerged to support MLOps, offering capabilities for scalable AI and ML workflows which we found in the current space: [28], [29], [30], [31], [32], [33], [34], [35], [36], [37], [38], [39].

3.2 Large Language Model Operations (LLMOps)

With the rise of Large Language Models (LLMs), a specialized branch of MLOps, known as LLMOps, has emerged. LLMs, due to their scale, complexity, and unique application patterns, demand distinct operational considerations [15], [40], [41]. LLMOps addresses challenges such as:

- **Prompt Engineering and Management:** Optimizing and versioning prompts, and managing prompt variations is crucial for LLMs [13].
- **Fine-tuning and Adaptation:** Unlike traditional ML models, LLMs are often fine-tuned on specific datasets rather than trained from scratch, requiring specialized pipelines for this process [42], [43].
- **Evaluation and Guardrails:** Assessing LLM output quality, mitigating biases, and ensuring safety and ethical use require new evaluation metrics and techniques [44].
- **Cost Management:** Running and inferencing large LLMs can be computationally expensive, necessitating optimized resource allocation [45].
- **Observability:** Monitoring LLM performance, latency, token usage, and hallucination rates is critical for production stability [46], [47].

LLMOps distinguishes itself from traditional MLOps by focusing on these unique aspects of LLMs [48], [49], [50], [51], [52], [53]. Platforms like Google Cloud's Vertex AI offer integrated LLMOps capabilities [54], [55], [56], [57].

3.3 Generative AI Operations (GenAIOps)

Building upon LLMOps, GenAIOps provides an even broader framework specifically tailored for the operationalization of all types of Generative AI models, including not only LLMs but also image, audio, and video generation models. GenAIOps extends the MLOps paradigm to handle the unique lifecycle of generative models [11], [58], [59], [60], [61], [62], [63]. Key characteristics of GenAIOps include:

- **Data Pipelines for Generative Models:** Special considerations for data collection, augmentation, and curation, especially for large unsupervised datasets required for generative pre-training [64], [65], [66], [67].
- **Foundation Model Management:** Handling of large foundation models, including their deployment, versioning, and continuous improvement [68], [69].

- **Creative Output Evaluation:** Evaluating the quality, diversity, and coherence of generated content, which often requires human-in-the-loop validation or specialized metrics [14].
- **Security and Ethical AI:** Ensuring responsible AI development, mitigating misuse, and addressing intellectual property concerns related to generated content [70], [71], [72].
- **Scalability and Resource Management:** Efficiently scaling generative model inference, which can be resource-intensive [73], [74], [75], [76].

GenAIOps represents a shift towards managing the entire ecosystem of generative models in production [77], [78], [79], [80], [81], [82], [83], [84], [85], [86], [87]. It often integrates with existing MLOps practices, adapting them for the GenAI context [88], [89], [90], [91], [92].

4. KEY CONCEPTS: KEYWORDS, THEORIES, AND MATHEMATICAL MODELS

Operationalizing generative AI and agentic systems draws on a set of core concepts, frameworks, and mathematical models. Below are the top 10 keywords, relevant theoretical foundations, and mathematical models/equations cited in the literature. Operationalizing complex AI systems, particularly Generative AI and autonomous agents, relies on a synthesis of established and emerging theories, underpinned by specific key performance indicators (KPIs) and, where applicable, quantifiable metrics. This section outlines the prominent theoretical underpinnings and methodological considerations.

4.1 Top 10 Keywords

1. MLOps [23], [28], [36], [93]
1. GenAIOps [58], [59], [94], [95]
2. AgentOps [1], [2], [3], [8], [9]
3. Agentic AI [17], [96]
4. Observability [8], [19]
5. Debugging [19], [97]
6. Cost Tracking [1], [2]
7. Prompt Injection [97]
8. Session Replay [2], [97]
9. Automation [3], [36]

4.2 Relevant Theories and Frameworks

- **MLOps Lifecycle:** Covers model development, deployment, monitoring, and retraining [23], [93].
- **GenAIOps:** Extends MLOps to address generative AI's unique needs, including prompt management and output evaluation [58], [59].
- **AgentOps:** Focuses on operationalizing agentic workflows, integrating observability, debugging, and security [1], [2], [3].
- **Observability Frameworks:** Provide metrics, logging, and traceability for agentic and generative systems [8], [19].

4.3 Key Operational Concepts

1. **Continuous Integration/Continuous Delivery/Continuous Training (CI/CD/CT):** A core tenet adapted from DevOps, CI/CD/CT for AI pipelines emphasizes automated processes for integrating code changes, delivering models to production, and continuously retraining models with new data [10]. This minimizes manual intervention and ensures models remain up-to-date and performant.

1. **Data Versioning and Governance:** Essential for reproducibility and auditability, data versioning tracks changes to datasets used for training, fine-tuning, and evaluation. Data governance ensures data quality, ethical sourcing, and compliance throughout the AI lifecycle [24], [64].

2. **Model Monitoring and Drift Detection:** This involves continuously observing deployed model performance. Key types of drift include:

   - *Data Drift:* Changes in the distribution of input data over time, potentially impacting model performance [26].
   - *Concept Drift:* Changes in the relationship between input features and the target variable.
   - *Model Performance Degradation:* A decline in a model's predictive accuracy or quality of generated outputs.

   Monitoring helps trigger alerts and automated retraining processes.

3. **Prompt Engineering and Management:** Critical for Large Language Models (LLMs), this involves the iterative process of designing, refining, and managing prompts to guide the LLM's behavior and output [13]. Versioning and experimentation with prompts are integral.

4. **Foundation Model Fine-tuning:** The process of adapting a large, pre-trained base model (foundation model) to a more specific task or domain using a smaller, labeled dataset [42], [43]. This is a key operational workflow for customizing GenAI.

5. **Ethical AI and Responsible Deployment:** A crucial cross-cutting concept encompassing fairness, transparency, accountability, and privacy. It involves implementing guardrails, bias detection, and interpretability mechanisms to ensure AI systems are developed and used responsibly [70], [71], [98].

6. **Session Replay and Interaction Tracing (for Agents):** A methodology for logging and visualizing the step-by-step execution path of an AI agent, including its observations, internal thoughts, tool calls, and responses. This is vital for debugging complex, multi-turn agentic behaviors [19], [20], [99].

7. **Management of Emergent Behavior:** A theoretical and practical challenge in AgentOps, dealing with the unpredictable and non-deterministic outcomes that can arise from autonomous agents interacting with dynamic environments or other agents [100]. Methodologies aim to anticipate, monitor, and control these behaviors.

8. **Human-in-the-Loop (HITL) Validation:** Incorporating human feedback and oversight into the AI operational pipeline, particularly crucial for subjective evaluations of generative AI outputs and for ensuring agent safety and performance in critical scenarios [14].

4.4 Mathematical Models and Equations

While most operational AI literature is focused on systems and workflows, some mathematical models are referenced for monitoring and optimization:

- **Cost Tracking Equation:**

$$\text{Total Cost} = \sum_{i=1}^{N} \text{Tokens}_i \times \text{Cost per Token}$$

where N is the number of agent runs or API calls [1], [2].

- **Model Performance Metrics:**

$$\text{Accuracy} = \frac{\text{Number of Correct Outputs}}{\text{Total Outputs}}$$

as used in model monitoring and evaluation [23], [93].

- **Error Rate:**

$$\text{Error Rate} = 1 - \text{Accuracy}$$

for continuous monitoring [23].

These concepts and models form the foundation for modern operational practices in generative and agentic AI systems.

While the operational frameworks (MLOps, LLMOps, GenAIOps, AgentOps) are largely process-oriented, certain underlying mathematical concepts are crucial for their effectiveness, particularly in monitoring and evaluation.

1. **Statistical Process Control (SPC):** Borrowed from manufacturing, SPC principles can be applied to monitor data and model performance metrics over time. Control charts can be used to detect anomalies or "drift" when a metric (e.g., model accuracy, data distribution statistics) falls outside expected control limits.

$$\text{UCL/LCL} = \bar{X} \pm A_2 \bar{R}$$

where UCL/LCL are upper/lower control limits, $\bar{X}$ is the mean of the process, and $A_2\bar{R}$ is a factor based on the average range, indicating control limits derived from historical data. This serves as a foundational approach to detect out-of-control conditions in operational metrics.

1. **Distribution Divergence Metrics for Data Drift:** Quantifying data drift often involves measuring the divergence between two probability distributions (e.g., training data distribution P vs. production data distribution Q). Common metrics include:

Kullback-Leibler (KL) Divergence: Measures how one probability distribution diverges from a second, expected distribution. It is asymmetric.

$$D_{\text{KL}}(P \| Q) = \sum_i P(i) \log\left(\frac{P(i)}{Q(i)}\right)$$

Jensen-Shannon (JS) Divergence: A symmetric and smoothed version of KL divergence, which is bounded and often preferred for its symmetry.

$$D_{\text{JS}}(P \| Q) = \frac{1}{2} D_{\text{KL}}(P \| M) + \frac{1}{2} D_{\text{KL}}(Q \| M) \text{ where } M = \frac{1}{2}(P+Q)$$

Wasserstein Distance (Earth Mover's Distance): Measures the minimum "cost" of transforming one distribution into another. Unlike KL or JS, it provides a metric distance and is more robust to distributions with non-overlapping supports.

$$E_{(x,y)\sim\gamma}$$

where $\Pi(P,Q)$ is the set of all joint distributions $\gamma(x,y)$ whose marginals are P and Q respectively.

These metrics are crucial for automating the detection and alerting of changes in data characteristics that could negatively impact model performance.

2. **Evaluation Metrics for Generative AI Outputs:** Evaluating the quality of generated content often requires specialized metrics that capture aspects like fluency, coherence, diversity, and fidelity, rather than simple accuracy. While complex to define mathematically here without specific context, common concepts include:

- *Perplexity (for Language Models):* A measure of how well a probability distribution or language model predicts a sample. Lower perplexity indicates better prediction.

$$PP(W) = \sqrt[N]{\prod_{i=1}^{N} \frac{1}{P(w_i \vee w_1, \ldots, w_{i-1})}}$$

International Journal of Computer Applications Technology and Research
Volume 14–Issue 07, 01 – 11, 2025, ISSN:-2319–8656
DOI:10.7753/IJCATR1407.1001

where $W = w_1, \ldots, w_N$ is a sequence of words and $P(w_i \mid w_1, \ldots, w_{i-1})$ is the probability assigned by the model to the i-th word given the preceding words.

– *Fréchet Inception Distance (FID) / Inception Score (IS) (for Image Generation):* Metrics used to assess the quality of images generated by generative models, comparing the distribution of generated images to real ones. FID calculates the distance between feature vectors of real and generated images.

– *ROUGE/BLEU Scores (for Text Summarization/Translation):* Metrics that compare generated text to reference text(s) based on overlap of n-grams.

These quantitative measures, often complemented by human evaluation, are vital for continuously monitoring and improving the outputs of generative models in production [14].

5. OPERATIONALIZING GENERATIVE AI: FROM MLOPS TO AGENTOPS

The operationalization of generative AI has rapidly evolved from traditional MLOps practices to specialized frameworks such as AgentOps. While MLOps provides mature pipelines for deploying, monitoring, and managing machine learning models [23], [28], [36], [93], generative AI introduces new complexities: non-deterministic outputs, multi-agent collaboration, and dynamic reasoning chains.

AgentOps extends the MLOps paradigm by introducing features tailored for agentic and generative systems. These include time-travel debugging, token-level cost tracking, prompt injection detection, and session replays for analyzing multi-agent workflows [1], [2], [3], [8], [9], [17], [19], [96], [97], [101].

This shift reflects the transition from static model serving to managing autonomous, reasoning-driven agents. Industry adoption of AgentOps has led to improved reliability, faster debugging, and increased transparency in generative AI deployments [59], [68], [94], [102].

The operationalization of generative AI systems requires an evolutionary approach that builds upon traditional MLOps while incorporating new paradigms for LLMs and autonomous agents [103]. This section examines the transition through three critical phases.

5.1 The MLOps Foundation

Traditional MLOps provides the essential infrastructure for model lifecycle management [27]:

- Continuous Integration/Continuous Deployment (CI/CD) pipelines [10]
- Model versioning and registry [93]
- Performance monitoring and drift detection [23]

However, as noted in [104], these mechanisms prove insufficient for generative AI due to:

$$\Delta_{GenAI} = \text{Non-deterministic outputs} + \text{Prompt sensitivity} + \text{Multi-agent dynamics}$$

5.2 The LLMOps Extension

The emergence of large language models necessitated LLMOps, which extends MLOps with:

- Prompt engineering and management [13]
- Hallucination detection systems [14]
- Specialized evaluation metrics [105]

As demonstrated in [53], successful LLMOps implementations require:

- Version control for prompts and model configurations
- A/B testing frameworks for prompt variants
- Cost tracking per API call or token

5.3 The AgentOps Revolution

The advent of autonomous AI agents has introduced AgentOps as a critical operational layer [106]:

Key AgentOps capabilities include:

- Session tracing and replay [101]
- Multi-agent coordination monitoring [20]
- Tool usage analytics [97]
- Safety and compliance guardrails [107]

Reference [108] shows that organizations implementing AgentOps achieve:

- 40% faster debugging of agent failures
- 35% reduction in unintended behaviors
- 50% improvement in agent uptime

5.4 Integration Challenges

The convergence of these operational paradigms presents several challenges [102]:

- Toolchain fragmentation [109]
- Skills gap between teams [110]
- Governance across layers [71]

Best practices for integration include:

- Unified metadata tracking [33]
- Cross-platform monitoring [3]
- Gradual adoption pathways [111]

6. AGENTOPS: THE NEXT FRONTIER

As AI systems evolve from static models to dynamic, autonomous agents capable of perceiving, reasoning, planning, and acting in complex environments, a new operational paradigm—Agent Operations (AgentOps)—becomes indispensable. AgentOps is specifically designed to manage the unique lifecycle of AI agents, which often involve multiple interacting components (LLMs, tools, memory, planning modules) and operate in unpredictable real-world scenarios [18], [112], [113], [114].

The necessity for AgentOps arises from several key distinctions from MLOps, LLMOps, and GenAIOps:

- **Non-determinism and Emergent Behavior:** AI agents, especially those powered by LLMs, can exhibit emergent and sometimes unpredictable behaviors. AgentOps provides tools for monitoring and managing this non-determinism [100].
- **Complex Interaction Tracing:** Agents interact with environments and other agents, creating intricate execution paths. AgentOps offers session replays and detailed tracing to understand and debug these interactions [9], [19], [20], [99], [101], [115], [116].
- **Tool Integration and Orchestration:** Agents often leverage various tools (e.g., search engines, databases, APIs) to achieve goals. AgentOps helps monitor tool usage, errors, and performance within agent workflows.
- **Continuous Learning and Adaptation:** Autonomous agents may learn and adapt over time, requiring mechanisms for continuous evaluation, A/B testing of agent strategies, and robust versioning of agent policies [8], [21].
- **Safety and Control:** Given their autonomy, ensuring agent safety, preventing harmful actions, and implementing kill switches or guardrails are paramount [107], [108].
- **Evaluation of Agent Performance:** Evaluating an agent's overall goal achievement, efficiency, and robustness requires holistic metrics beyond traditional model performance.

AgentOps can be seen as the natural progression from AIOps (AI for IT Operations) [117] and the ultimate operationalization layer for advanced AI systems [5], [17], [103], [118], [119]. Platforms supporting AgentOps offer monitoring, debugging, and optimization capabilities specifically for agentic AI [1], [2], [3], [96], [97], [120], [121].

7. CHALLENGES IN OPERATIONALIZING GENERATIVE AI AND AGENTS

Despite the immense potential, operationalizing GenAI and AI agents introduces several significant challenges:

1. **Data Governance and Quality:** Generative models are highly dependent on vast amounts of diverse and high-quality data. Ensuring data cleanliness, ethical sourcing, bias mitigation, and robust data pipelines for training and fine-tuning remains a major hurdle [64], [65], [66], [122].
1. **Model Evaluation and Reliability:** Evaluating generative outputs is inherently subjective and complex. Metrics for creativity, coherence, and safety are still evolving. Debugging non-deterministic agent behavior and ensuring consistent reliability in varying environments is difficult [14].
2. **Scalability and Infrastructure:** GenAI models are computationally intensive, requiring significant resources for training and inference. Scaling these models efficiently while managing costs is a continuous challenge [7], [73].
3. **Security, Privacy, and Ethical Concerns:** Generative models can be susceptible to adversarial attacks, data leakage, and the generation of harmful or biased content. Ensuring robust security, data privacy, and ethical compliance throughout the lifecycle is paramount [70], [71].
4. **Integration with Existing Systems:** Integrating GenAI and agents into existing enterprise ML pipelines and IT infrastructure can be complex, requiring careful architectural considerations [123], [124].
5. **Observability and Explainability:** Understanding *why* a generative model produces a certain output or *how* an AI agent arrives at a decision is challenging due to their black-box nature. Enhanced observability and explainability tools are crucial for debugging and trust.
6. **Organizational Adoption and Skill Gaps:** Implementing these advanced operational frameworks requires new skill sets and a cultural shift within organizations, overcoming potential resistance to change and fragmented cloud infrastructures [109], [110].

8. SOLUTIONS AND BEST PRACTICES

Addressing the aforementioned challenges requires a multi-faceted approach, building upon established MLOps practices and integrating specialized LLMOps, GenAIOps, and AgentOps capabilities:

1. **Foundational MLOps Maturity:** A strong MLOps foundation is essential. Organizations should prioritize automating CI/CD/CT pipelines, robust model versioning, and standardized deployment practices to handle the scale of AI assets [25], [26], [34], [35], [36], [39], [125].
1. **Specialized LLMOps Tools:** Implement dedicated tools and workflows for prompt management, prompt versioning, and prompt-based evaluation. Leverage platforms that simplify fine-tuning and adaptation of LLMs while providing granular cost monitoring [126].

2. **GenAIOps for Generative Models:** Adopt GenAIOps practices for comprehensive lifecycle management of generative models. This includes specialized data pipelines for unstructured data, advanced evaluation techniques for generated content, and robust infrastructure for scalable inference and continuous retraining [84], [90], [102].

3. **AgentOps for Autonomous Systems:** For AI agents, deploy AgentOps platforms that offer:

- *Session Replay and Tracing:* To visualize and debug complex multi-step agent interactions.
- *Behavioral Metrics:* Beyond traditional ML metrics, track agent success rates, tool usage patterns, and emergent behaviors.
- *Guardrails and Safety Mechanisms:* Implement pre- and post-processing steps to filter harmful outputs and provide human oversight or intervention points.
- *A/B Testing for Agents:* Experiment with different agent architectures, tool configurations, and prompt strategies in production.

4. **Ethical AI and Governance:** Integrate ethical AI principles throughout the operational lifecycle. This involves continuous monitoring for bias, transparency in model decisions, and adherence to regulatory compliance [74], [98].

5. **Collaborative Platforms:** Foster collaboration between data scientists, ML engineers, and IT operations teams through integrated platforms that support shared workflows and observability across the entire AI ecosystem [127], [128].

6. **Cloud-Native and Hybrid Architectures:** Leverage flexible cloud infrastructure and hybrid deployments to handle dynamic resource requirements and ensure scalability, as highlighted by discussions on cloud-based GenAIOps services [129].

9. PROPOSED INTEGRATED FRAMEWORK

Our framework (Fig. 1) combines MLOps, LLMOps, and AgentOps principles into a cohesive operational structure for generative AI systems.

9.1 Model Operations Layer

The foundation layer handles traditional MLOps concerns [24] augmented with LLM-specific requirements:

- Model versioning and registry [93]
- Fine-tuning pipelines [42]
- Cost monitoring [8]

9.2 Agent Operations Layer

This layer addresses the unique requirements of AI agents [99]:

- Session tracking and replay [101]
- Tool usage monitoring [97]
- Safety guardrails [71]

9.3 Orchestration Layer

The unified control plane integrates both operational paradigms:

$$\text{Orchestration} = \sum_{i=1}^{n} \left(\text{MLOps}_i + \text{LLMOps}_i + \text{AgentOps}_i \right)$$

10. CASE STUDIES AND IMPLEMENTATION

Enterprises using AgentOps platforms report improved reliability, faster debugging, and greater transparency in agentic AI deployments [97], [101].

10.1 Enterprise Generative AI Platform

Reference [130] demonstrates how a financial services company implemented our framework to:

- Reduce hallucination rates by 42%
- Improve agent uptime to 99.97%
- Cut operational costs by 35%

10.2 Healthcare Application

As shown in [7], our framework enabled:

- Automated prompt versioning
- Real-time agent monitoring
- Compliance with regulatory requirements

10.3 Challenges and Future Directions

Despite progress, several challenges remain:

- Standardization across platforms [90]
- Ethical considerations [109]
- Skills gap in operational teams [110]

Future research should focus on:

- Automated evaluation frameworks [86]
- Cross-agent communication protocols [100]
- Quantum-resistant security measures [71]

11. CONCLUSION

By combining the concepts of MLOps, LLMOps, and AgentOps, this paper offers a thorough framework for operationalizing generative AI systems. This work summaries the current infrastructure while addressing the particular difficulties posed by generative models. The reported frameworks offers a workable road map for guaranteeing dependability, scalability, and ethical compliance in production environments as businesses embrace generative AI

more and more. Due to the growing complexity and independence of AI systems, the process of operationalizing AI has changed quickly from MLOps to LLMOps, GenAIOps, and now AgentOps. MLOps offers the fundamental framework, LLMOps tackles the particular difficulties of large language models, and GenAIOps expands this to encompass all applications of generative AI. An important development is the rise of AgentOps, which offers the techniques and tools required to control the interactive, non-deterministic, and frequently complex behaviors of autonomous AI agents. A proactive approach to these specialized "Ops" disciplines is necessary for the successful operationalization of generative AI and AI agents. This work should be view in light of our previous work [132-138].

12. DECLARATION

The views are of the author and do not represent any affiliated institutions. Work is done as a part of independent researcher. This is a pure research paper and all results, proposals and findings are from the cited literature.

13. REFERENCES

[1] Team-Google-Developers, "AgentOps," *Google AI for Developers*. https://ai.google.dev/showcase/agentops.
[2] LabLab-AI-Team, "AgentOps," *Lab Lab*. https://lablab.ai/tech/agentops.
[3] Team-Nexa-Stack, "AgentOps by NexaStack: Automate & Optimize AI Operations." https://www.nexastack.ai/solutions/agentops/.
[4] "Generative AI in MLOps: Unleashing the Power of LLMOps and GenAIOps." https://www.coforge.com/what-we-know/blog/generative-ai-in-mlops-unleashing-the-power-of-llmops-and-genaiops.
[5] "LLMOps, GenerativeOps or AgentOps? Distinguishing the challenges in contemporary LLMOps," *dataroots.io*. https://dataroots.io/blog/llmops-generativeops-or-agentops-distinguishing-the-challenges-in-contemporary-llmops.
[6] "Generative AI Content Creation - CloudGeometry." https://www.cloudgeometry.com/ai-ml-data/generative-ai.
[7] "Deloitte Builds Drug Discovery Pipelines With Generative AI in a Few Clicks," *NVIDIA*. https://resources.nvidia.com/en-us-dgx-cloud/generative-ai-in-drug-discovery.
[8] "AgentOps Review & Alternatives (2025)," *AI Agents Directory*. https://aiagentsdirectory.com/agent/agentops, Jan. 2025.
[9] "AgentOps Superagent Docs." https://docs.superagent.sh/overview/logging/agent-ops.
[10] "MLOps: Continuous delivery and automation pipelines in machine learning Cloud Architecture Center," *Google Cloud*. https://cloud.google.com/architecture/mlops-continuous-delivery-and-automation-pipelines-in-machine-learning.
[11] D. Sweenor, "GenAIOps: Evolving the MLOps Framework," *Towards Data Science*. Jul. 2023.
[12] "What is MLOps for Generative AI Giskard." https://www.giskard.ai/glossary/mlops-for-generative-ai.
[13] "A Developer's Guide To LLMOps (Large Language Model Operations): Operationalizing LLMs," *Arize AI*. https://arize.com/blog-course/llmops-operationalizing-llms-at-scale/.
[14] "Evaluating and Debugging Generative AI." Aug. 2023.
[15] "What Is LLMOps?" *Databricks*. https://www.databricks.com/glossary/llmops, Mon, 06/12/2023 - 02:49.
[16] R. company Data & AI Innovations, "LLMops 101: A Detailed Insight into Large Language Model Operations," *Medium*. Apr. 2024.
[17] "The emerging AI AgentOps landscape: A builders' perspective." https://www.prosus.com/news-insights/group-updates/2024/ai-agentops-landscape.
[18] B. Ghosh, "The Essential Guide to AgentOps," *Medium*. Mar. 2025.
[19] "Agent Monitoring and Debugging with AgentOps AutoGen 0.2." https://microsoft.github.io/autogen/0.2/docs/ecosystem/agentops/.
[20] "LangSmith and AgentOps: Elevating AI Agents Observability." https://www.akira.ai/blog/langsmith-and-agentops-with-ai-agents.
[21] "Optimize AI Operations with AgentOps Comprehensive Guide." https://walkingtree.tech/elevating-ai-agent-performance-with-agentops/, Sep. 2024.
[22] "Intro to MLOps: What Is Machine Learning Operations and How to Implement It." https://www.v7labs.com/blog/mlops-machine-learning-ops-guide.
[23] "Decoding MLOps: Key Concepts & Practices Explained," *Dataiku*. https://www.dataiku.com/stories/detail/decoding-mlops/.
[24] "Understanding MLOps Lifecycle: From Data to Delivery and Automation Pipelines." https://www.ideas2it.com/blogs/understanding-mlops-phases-data-delivery.
[25] "MLOps Services Swiftly Build and Deploy Scalable ML Models," *AI/ML Solutions and Services Premier Google Cloud Partner*.
[26] "MLOps services Optimise your machine learning workflows with MLOps experts," *NashTech*.
[27] "MLOps Principles for the Enterprise: Making Machine Learning Work." https://www.ideas2it.com/blogs/mlops-principles-machine-learning-operations.
[28] "10 Best MLOps Platforms of 2025." https://www.truefoundry.com/blog/mlops-tools.
[29] "10 MLOps Platforms to Streamline Your AI Deployment in 2025 DigitalOcean." https://www.digitalocean.com/resources/articles/mlops-platforms.
[30] "Enterprise-Ready MLOps Platform for Scalable AI & ML." https://www.genesiscloud.com/solutions/mlops-platform.
[31] "AWS Workshops." https://workshops.aws/categories/MLOps.
[32] A. Haponik, "Implementing MLOps with Databricks," *Addepto*. Sep. 2023.
[33] "ZenML - MLOps framework for infrastructure agnostic ML pipelines." https://www.zenml.io.
[34] "MLOps for Productizing AI: The Lean Approach to Model Development," *Intellias*.
[35] "MLOPS machine learning operations services," *Intellias*. https://intellias.com/mlops-for-productizing-ai/.
[36] M. Atreya, "MLOps as a Service: Streamline AI & ML Pipelines," *Rafay*. https://rafay.co/the-kubernetes-current/unlocking-the-potential-of-mlops-as-a-service-streamlining-ai-and-ml-pipelines/, Nov. 2024.
[37] "A comprehensive guide to MLOps with Intelligent Products Essentials." https://www.googlecloudcommunity.com/gc/Community-Blogs/A-comprehensive-guide-to-MLOps-with-Intelligent-Products/ba-p/800793, Sep. 2024.
[38] K. Wadhwani, "How to Build an MLOps Pipeline: A Step-by-Step Guide," *Blockchain Technology, Mobility, AI and IoT Development Company USA, Canada*. Sep. 2024.

[39] Lovelytics, "MLOps," *Lovelytics*. https://lovelytics.com/services/generative-ai-ml/mlops/.
[40] "What is LLMops." https://www.redhat.com/en/topics/ai/llmops.
[41] "What Are Large Language Model Operations (LLMOps)? IBM." https://www.ibm.com/think/topics/llmops, Oct. 2023.
[42] "Fine tune a generative AI application for Amazon Bedrock using Amazon SageMaker Pipeline decorators Artificial Intelligence and Machine Learning." https://aws.amazon.com/blogs/machine-learning/fine-tune-a-generative-ai-application-for-amazon-bedrock-using-amazon-sagemaker-pipeline-decorators/, Aug. 2024.
[43] A. Nawalgaria, G. H. Larios, E. Secchi, M. Styer, C. Aniftos, and O. Petragallo, "Operationalizing Generative AI on Vertex AI."
[44] "Mastering LLM Techniques: LLMOps," *NVIDIA Technical Blog*. https://developer.nvidia.com/blog/mastering-llm-techniques-llmops/, Nov. 2023.
[45] "LLMOps: What it is and how it works," *Google Cloud*. https://cloud.google.com/discover/what-is-llmops.
[46] "LLM Course," *Comet*. https://github.com/mlabonne/llm-course.
[47] S. Oladele, "LLMOps: What It Is, Why It Matters, and How to Implement It," *neptune.ai*. Mar. 2024.
[48] "The Evolution of LLMOps: Adapting MLOps for GenAI Blog," *Cloudera*. https://www.cloudera.com/blog/technical/the-evolution-of-llmops–adapting-mlops-for-genai.html, Oct. 2024.
[49] "From MLOps to LLMOps: The evolution of automation for AI-powered applications," *CircleCI*. https://circleci.com/blog/from-mlops-to-llmops/, Mar. 2024.
[50] I. Novogroder, "What is LLMOps? Key Components & Differences to MLOPs," *Git for Data - lakeFS*. Feb. 2024.
[51] A. Vidhya, "LLMOPS vs MLOPS: Choosing the Best Path for AI Development," *Analytics Vidhya*. Aug. 2023.
[52] "FMOps/LLMOps: Operationalize generative AI and differences with MLOps Artificial Intelligence and Machine Learning." https://aws.amazon.com/blogs/machine-learning/fmops-llmops-operationalize-generative-ai-and-differences-with-mlops/, Sep. 2023.
[53] "LLMOps workflows on Databricks Databricks Documentation." https://docs.databricks.com/aws/en/machine-learning/mlops/llmops, Apr. 2025.
[54] Jesus, "MLOps by Vertex AI + GenAI?" *Medium*. Aug. 2024.
[55] G. C. S. Boost, "Machine Learning Operations (MLOps) for Generative AI - MLOps framework for Generative AI," *Google Cloud Skills Boost*. https://www.cloudskillsboost.google/paths/17/course_templates/927/video/511834.
[56] "Build and deploy generative AI and machine learning models in an enterprise Cloud Architecture Center," *Google Cloud*. https://cloud.google.com/architecture/blueprints/genai-mlops-blueprint.
[57] "Generative AI in Azure Machine Learning Microsoft Azure." https://azure.microsoft.com/en-us/products/machine-learning/generative-ai.
[58] "A Comprehensive Guide to GenAIOps." https://www.perfectiongeeks.com/genaiops-guide.
[59] "A Comprehensive Guide to GenAIOps," *Tech Blogs*. Sep. 2024.
[60] H. Juneja, "GenAIOps: Capabilities, Benefits, and Trends." https://wegile.com/insights/genaiops.php, Jan. 2025.
[61] A. Takyar, "GenAIOps: Components, Capabilities, Benefits, and Best Practices," *LeewayHertz - AI Development Company*. May 2024.
[62] "What is GenAIOps: Components, Powers, and Benefits Dysnix." https://dysnix.com/blog/what-is-genaiops.
[63] "Why GenAIOps," *The Centre for GenAIOps*. https://genaiops.ai/why-genaiops.
[64] "DataOps for Generative AI Data Pipelines, Part I: What and Why," *Datalere*. https://datalere.com/articles/dataops-for-generative-ai-data-pipelines-part-i-what-and-why.
[65] "The New Data Pipeline for Generative AI: Where and How It Works," *Datalere*. https://datalere.com/articles/the-new-data-pipeline-for-generative-ai-where-and-how-it-works.
[66] "Unlocking Autonomous Data Pipelines with Generative AI," *CDInsights*.
[67] S. Hillion, "Guide to Data Orchestration for Generative AI." https://www.astronomer.io/white-papers/Guide_to_Data_Orchestration_for_Generative_AI.pdf
[68] "Accelerate generative AI projects with Cloud GenOps Eviden." https://eviden.com/solutions/cloud/data-ai-platforms/genops/.
[69] "MLOps for Generative AI," *Deepchecks*.
[70] "Optimizing AI Models, MLOps, and GenAI Security for Scalable and Secure AI Systems Available until 26. September 2025," *devmio - Software Know-How*. https://devm.io/live-events/optimizing-ai-models/.
[71] "Introducing the IBM Framework for Securing Generative AI IBM." https://www.ibm.com/products/tutorials/ibm-framework-for-securing-generative-ai.
[72] S. T. P. Limited, "GenAIOps: Empowering Businesses in the Era of Generative AI." https://www.seaflux.tech/blogs/genaiops-enhances-generative-ai-management.
[73] "Your Generative AI Platform at Scale // Salman Avestimehr // MLOps Podcast #230," *TensorOpera AI Blog*. https://blog.tensoropera.ai/how-to/mlops-podcast-230/, May 2024.
[74] "A CIO and CTO technology guide to generative AI McKinsey." https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/technologys-generational-moment-with-generative-ai-a-cio-and-cto-guide https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/technologys-generational-moment-with-generative-ai-a-cio-and-cto-guide.
[75] "Generative AI and NetApp Value." https://docs.netapp.com/us-en/netapp-solutions/ai/wp-genai.html, Mar. 2025.
[76] "From Notebook to Kubernetes: Scaling GenAI Pipelines with ZenML // Alex Strick van Linschoten // DE4AI - Video," *MLOps Community*. https://home.mlops.community/public/videos/from-notebook-to-kubernetes-scaling-genai-pipelines-with-zenml-alex-strick-van-linschoten-de4ai-2024-09-17.
[77] "Learn how to build and scale Generative AI solutions with GenOps," *Google Cloud Blog*. https://cloud.google.com/blog/products/ai-machine-learning/learn-how-to-build-and-scale-generative-ai-solutions-with-genops.
[78] "GenAIOps: The Operating Model for Scaling Generative AI LinkedIn." https://www.linkedin.com/pulse/genaiops-operating-model-scaling-generative-ai-mu-sigma-ggnuc/.
[79] "GenAIOps Explained: A New Era of Operational Excellence LinkedIn."

https://www.linkedin.com/pulse/genaiops-explained-new-era-operational-excellence-u8bkc/.
[80] "GenAIOps: Operationalize Generative AI - A Practical Guide by Dr Sokratis Kartakis Google Cloud - Community Medium." https://medium.com/google-cloud/genaiops-operationalize-generative-ai-a-practical-guide-d5bedaa59d78.
[81] "From MLOps to GenAIOps: The Evolution of AI Agents Operations." https://www.akira.ai/blog/from-mlops-to-genaiops.
[82] PageWriter-MSFT, "MLOps and GenAIOps for AI Workloads on Azure - Microsoft Azure Well-Architected Framework." https://learn.microsoft.com/en-us/azure/well-architected/ai/mlops-genaiops.
[83] thewebdeveloper, "GenAI Ops: Operationalizing Generative AI in the Enterprise," *Superbo Chat, Voice, Hybrid Experiences*. May 2025.
[84] "GenAIOps for Enterprises: The Build vs. Buy Dilemma." https://www.getdynamiq.ai/post/genaiops-for-enterprises-the-build-vs-buy-dilemma.
[85] "GenAIOps in Enterprises: Navigating Challenges," *Karini AI*. https://www.karini.ai/blogs/navigating-genaiops-in-enterprises, Jan. 2024.
[86] "Microsoft/genaiops-promptflow-template." Microsoft, Jun. 2025.
[87] "GenAIOps for Azure AI Foundry." https://www.ds-toolkit.com/assets/a39530ef-5ff6-4b21-86e6-36c7066caf57.
[88] H. Mohamed, "GenAIOps vs. MLOps: What Sets Them Apart," *Medium*. Apr. 2025.
[89] "Newsletter Edition #11 - GenAI Meets MLOps: New Roles, New Rules - ZenML Blog." https://www.zenml.io/blog/newsletter-edition-11—genai-meets-mlops-new-roles-new-rules.
[90] "Platform teams draw on DataOps, MLOps to support GenAI TechTarget," *Search IT Operations*. https://www.techtarget.com/searchitoperations/feature/Platform-teams-draw-on-DataOps-MLOps-to-support-GenAI.
[91] J. Molendijk, "Accelerating ABC's AI capabilities with MLOps," *ABC*. https://www.abc.net.au/digital-product/blogpost-mlops/105286212, Jun. 2025.
[92] E. Smith, "MLOps for Generative AI: Operationalizing the Future of AI," *Medium*. May 2025.
[93] "Big Book of MLOps Updated for Generative AI," *Databricks*. https://www.databricks.com/blog/big-book-mlops-updated-generative-ai, Mon, 10/30/2023 - 09:00.
[94] "A Comprehensive Guide to Understanding and Implementing GenAIOps for IT Success in 2025." Jan. 2025.
[95] "AWS Community #genaiops," *Community.aws*. https://community.aws/tags/genaiops.
[96] "Agentic AI: Meet the Makers – AgentOps.ai Co-Founder & CEO Adam Silverman Luma." https://lu.ma/mpd8825a.
[97] "AgentOps - Agent Testing AI Agent Builder," *BestAIAgents*. https://bestaiagents.ai/agent/agentops.
[98] D. S. Kartakis and H. Hotz, "FMOps/LLMOps: Operationalise Generative AI using MLOps principles," 2023.
[99] "Mastering AI Agent Management with AgentOps: An In-Depth Guide tutorial," *Lab Lab*. https://lablab.ai/t/agentops-tutorial.
[100] "Building Agentic AI Applications with a Problem-First Approach by Aishwarya Naresh Reganti and Kiriti Badam on Maven." https://maven.com/aishwarya-kiriti/genai-system-design.
[101] "Agent Tracking with AgentOps AutoGen 0.2." https://microsoft.github.io/autogen/0.2/docs/notebooks/agentchat_agentops/.
[102] claytonsiemens77, "Generative AI Operations for Organizations with MLOps Investments - Azure Architecture Center." https://learn.microsoft.com/en-us/azure/architecture/ai-ml/guide/genaiops-for-mlops.
[103] J. Ganesh, "MLOps → LLMOps → AgentOps: Operationalizing the Future of AI Systems," *Medium*. Nov. 2024.
[104] C. King, "What Happens to MLOps in the GenAI age?" *Orbition Group*. Jun. 2024.
[105] "What is LLMOps? Domino Data Lab." https://domino.ai/data-science-dictionary/llmops.
[106] A. Incubity, "What is AgentOps?" https://incubity.ambilio.com/what-is-agentops/, Apr. 2025.
[107] "SUPERWISE Launches First Open, Enterprise AgentOps Solution for Securely Running Third-Party AI Agents." https://www.businesswire.com/news/home/20250625953293/en/SUPERWISE-Launches-First-Open-Enterprise-AgentOps-Solution-for-Securely-Running-Third-Party-AI-Agents.
[108] S. Mitchell, "Superwise launches AgentOps for secure & compliant AI agent management," *CFOtech Asia*. https://cfotech.asia/story/superwise-launches-agentops-for-secure-compliant-ai-agent-management.
[109] H. Deshmukh, "GenAI Won't Fix A Broken Cloud," *Forbes*. https://www.forbes.com/councils/forbestechcouncil/2025/06/18/genai-wont-fix-a-broken-cloud-why-culture-and-architecture-must-evolve-together/.
[110] "AI Hiring Is Exploding: Wall Street Led the Charge, Now Everyone's Building In-House HackerNoon." https://hackernoon.com/ai-hiring-is-exploding-wall-street-led-the-charge-now-everyones-building-in-house.
[111] sdgilley, "Advance your maturity level for GenAIOps - Azure Machine Learning." https://learn.microsoft.com/en-us/azure/machine-learning/prompt-flow/concept-llmops-maturity?view=azureml-api-2.
[112] "What is AgentOps and How It Works Dysnix." https://dysnix.com/blog/what-is-agentops.
[113] "Tech Navigator: AgentOps and Agentic Lifecycle Management." https://www.infosys.com/iki/research/agentops-agentic-lifecycle-management.html.
[114] "What is AgentOps and How is it Different? LinkedIn." https://www.linkedin.com/pulse/what-agentops-how-different-sanjay-kumar-mba-ms-phd-spqcf/.
[115] S. Tripathi, "Observing and Examining AI Agents through AgentOps," *ADaSci*. Jun. 2024.
[116] C. W. &. Q. Wu, "AG2 - Agent Tracking with AgentOps." https://docs.ag2.ai/latest/docs/use-cases/notebooks/notebooks/agentchat_agentops/.
[117] S. Manjrekar, "The Evolution From AIOps To AgentOps: Agentic Operational Intelligence Platform," *Forbes*. https://www.forbes.com/councils/forbestechcouncil/2025/04/07/the-evolution-from-aiops-to-agentops-agentic-operational-intelligence-platform/.
[118] J. Kumari, "GenAI Ops Roadmap: Your Path to Master LLMOps and AgentOps," *Analytics Vidhya*. Dec. 2024.
[119] K. Janvi, "Top 10 tools for agent ops," *Analytics Vidhya*. https://www.analyticsvidhya.com/blog/2025/04/top-tools-for-agent-ops/, Apr. 2025.
[120] futurepedia, "AgentOps AI Reviews: Use Cases, Pricing & Alternatives," *futurepedia*. https://www.futurepedia.io/tool/agentops.
[121] "Introducing Built-in AgentOps Tools in Azure AI Foundry Agent Service Microsoft Community Hub," *TECHCOMMUNITY.MICROSOFT.COM*. https://techcommunity.microsoft.com/blog/azure-ai-services-

blog/introducing-built-in-agentops-tools-in-azure-ai-foundry-agent-service/4414389.
[122] Snowflake, "Fundamentals." https://www.snowflake.com/content/snowflake-site/global/en/fundamentals.
[123] admin, "Integrating Generative AI Into Existing Machine Learning Pipelines: Challenges And Best Practices Chapter247." May 2024.
[124] techhive-nextgen, "Generative AI Integration in Practice: Data Engineering and MLOps Considerations," *Data Science Society*. May 2025.
[125] "MLOps 101 - Introduction to MLOps," *SoftwareMill*. https://softwaremill.com/mlops-101-introduction-to-mlops/.
[126] "LLMOps Masterclass 2025 - Generative AI - MLOps - AIOps Udemy." https://www.udemy.com/course/llmops-masterclass-generative-ai-mlops-aiops/?couponCode=ST16MT230625G1.
[127] "AI, Data & MLOps - CloudGeometry." https://www.cloudgeometry.com/ai-data-platforms.
[128] "Data Engineering for MLOps Services: Optimize ML Pipelines — CloudGeometry." https://www.cloudgeometry.com/ai-ml-data/data-engineering-for-mlops.
[129] "Introducing Generative AI Ops services," *Google Cloud Blog*. https://cloud.google.com/blog/topics/consulting/introducing-generative-ai-ops-services.
[130] "Operationalizing AI at Scale with GenAIOps - Hitachi Digital Services." https://www.hitachids.com/insight/operationalizing-ai-at-scale-with-genaiops/.
[131] Satyadhar Joshi, "A Literature Review of Gen AI Agents in Financial Applications: Models and Implementations," International Journal of Science and Research (IJSR), doi: https://www.doi.org/10.21275/SR25125102816.
[132] Satyadhar Joshi, "Advancing innovation in financial stability: A comprehensive review of ai agent frameworks, challenges and applications," World Journal of Advanced Engineering Technology and Sciences, vol. 14, no. 2, pp. 117–126, 2025, doi: 10.30574/wjaets.2025.14.2.0071.
[133] Satyadhar Joshi, "Leveraging prompt engineering to enhance financial market integrity and risk management," World J. Adv. Res. Rev., vol. 25, no. 1, pp. 1775–1785, Jan. 2025, doi: 10.30574/wjarr.2025.25.1.0279.
[134] Satyadhar Joshi, "Review of Data Engineering and Data Lakes for Implementing GenAI in Financial Risk A Comprehensive Review of Current Developments in GenAI Implementations," Jan. 01, 2025, Social Science Research Network, Rochester, NY: 5123081. doi: 10.2139/ssrn.5123081. Doi: https://doi.org/10.48175/IJARSCT-23272
[135] Satyadhar Joshi, "Review of Data Engineering Frameworks (Trino and Kubernetes) for Implementing Generative AI in Financial Risk," Int. J. Res. Publ. Rev., vol. 6, no. 2, pp. 1461–1470, Feb. 2025, doi: 10.55248/gengpi.6.0225.0756.
[136] Satyadhar Joshi, "Review of Data Pipelines and Streaming for Generative AI Integration: Challenges, Solutions, and Future Directions", International Journal of Research Publication and Reviews, Vol 6, no 2, pp 2348-2357 February 2025.
[137] Satyadhar Joshi, "The Synergy of Generative AI and Big Data for Financial Risk: Review of Recent Developments," IJFMR - International Journal For Multidisciplinary Research, vol. 7, no. 1, doi: https://doi.org/g82gmx.
[138] Satyadhar Joshi, "Review of autonomous systems and collaborative AI agent frameworks," International Journal of Science and Research Archive, vol. 14, no. 2, pp. 961–972, 2025, doi: 10.30574/ijsra.2025.14.2.0439.