Spectrum of Engineering Sciences
ISSN (e) 3007-3138 (p) 3007-312X Volume 3, Issue 7, 2025

EXPLAINABLE AIOPS: A DEEP SURVEY ON TRUSTWORTHY AND TRANSPARENT AI IN CLOUD-SCALE DEVOPS AUTOMATION

Muhammad Adnan Sami[1], Abdur Rehman[2], Zahoor Ahmad[3], Nudrat Bano[*4]

[1,3,4]*Department of computer science, University of Agriculture, Faisalabad, Pakistan*
[2]*Department of Software Engineering, University of Malakand Chakdara Dir Lower KP*

[1]adnansamii.pk@gmail.com, [2]abdurrehman@uom.edu.pk, [3]i.am.zahoor.a@outlook.com[2],
[*4]nudrataliahmad@gmail.com

DOI: https://doi.org/10.5281/zenodo.15867960

Keywords

Article History
Received: 04 April, 2025
Accepted: 25 June, 2025
Published: 12 July, 2025

Corresponding Author: *
Nudrat Bano

Abstract

As Artificial Intelligence for IT Operations (AIOps) becomes increasingly integral to managing modern, cloud-scale DevOps environments, concerns about the opacity of AI-driven decisions have grown significantly. The adoption of black-box models in these systems, while enabling rapid automation, introduces critical challenges in trust, auditability, and regulatory compliance. This undermines confidence in automated decisions, especially in high-stakes environments involving anomaly detection, root cause analysis, and predictive scaling.

To address these challenges, the integration of Explainable Artificial Intelligence (XAI) into AIOps has emerged as a key research and industrial imperative. XAI techniques promise to make AI behaviors transparent, interpretable, and human-understandable, thus facilitating trust and control in AI-augmented operational workflows. However, the landscape of explainable AIOps is fragmented across tools, methods, and deployment scenarios, and there is no consolidated resource that comprehensively maps the field.

This survey presents a deep, structured analysis of over 70 scholarly works spanning academic literature and industry applications. We develop a taxonomy of explainable AIOps techniques, categorizing them by method type (e.g., model-agnostic, deep learning-based, symbolic), DevOps use-case (e.g., monitoring, RCA, auto-remediation), and cloud-specific integration. Through this analysis, we highlight significant gaps in scalability, standardization, and usability that current approaches fail to address in dynamic cloud environments.

Finally, the paper provides a roadmap for future research directions, including hybrid neuro-symbolic explainability, human-in-the-loop systems, and edge-cloud trustworthy AIOps. This work aims to serve as a foundational reference for researchers and practitioners seeking to build transparent, trustworthy, and scalable AI systems for modern cloud operations.

INTRODUCTION

In recent years, Artificial Intelligence for IT Operations (AIOps) has rapidly transformed how large-scale cloud-based systems are managed, monitored, and optimized. As modern DevOps environments grow increasingly complex—incorporating micro services, containerization, continuous deployment, and multi-cloud infrastructures—the manual handling of monitoring,

fault detection, and response becomes both inefficient and risky. AIOps works around this by automating the decision-making process along the IT pipeline through artificial intelligence and machine learning. Nevertheless, fast and large scale of AIOps is undeniable, but its transparency is compromised by the use of black-box models, which presents a major impediment to trust in high-stakes enterprise applications where explainability is essential to regulatory compliance, debugging, and accountability (Lyu et al., 2021).

The very nature of the AIOps models being uninterpretable causes the main problem because in most cases, they do not allow seeing why an action was predicted, why an anomaly occurred, or what a remediation action was initiated. This unclear nature becomes especially concerning in production spaces where teams on operations are liable to all the actions made by automated systems. The growth of explainable AI (XAI) in DevOps requires researchers and practitioners to study the framework that could reveal the internal logic of AIOps systems, so that the performance and real-time efficiency remain. In that regard, explainability is not only an improvement that makes changes in AIOps systems more readable, but rather a pillar of creating trustworthy AIOps ecosystems (Garg et al., 2023).

To solve the explainability gap, scientists have proposed a collection of XAI techniques such as model-agnostic (SHAP and LIME), attention-based frameworks, which display how a model forms a decision on a time-series dataset. These tools, however, must be adapted to dynamic DevOps pipelines, which is potentially challenging, including when it comes to processing high-volume log data, asynchronous events, and when it comes to enabling real-time decision-making at scale. Although the interest is increasing there is still a deficit of unified frameworks or taxonomies that would consolidate XAI techniques with a focus on an AIOps environment (Nguyen et al., 2023). Current solutions are frequently very narrowly devoted to individual modules e.g. anomaly detection or log parsing with little to no end to end visibility of the DevOps lifecycle. Moreover, explainability tools performing well in a static setting tend to struggle in a dynamic cloud-native setting where both configurations, workloads, and failure patterns are constantly changing (Trivedi et al., 2024).

This segregation means that there are major usability and adoption obstacles. It is not always straightforward to integrate explainability into continuous delivery pipeline or incident response process as DevOps teams frequently find it challenging. The explanations in real life applications are technically complex, therefore, operators cannot understand them, or too ambiguous, where they cannot make decisions. Also, explainable AIOps should support performance and be able to estimate that performance is not negatively adjusted by adding computation of additional explanation layers that providers/components of the system exhibit in response to system-wide outages or when important degradation in those providers/components occurs (Khalasi et al., 2023). The problems present a burning question of standardization, benchmarking and improved design of incorporating XAI in the AIOps tools chain.

In the very light of these shortcomings, this paper undertakes a complete survey of the explainable AIOps by evaluating more than 70 scholarly and practice-oriented research papers. Overall, its aim is to put together the disparate state of prior art and present an organized survey of methods, instruments, and implementations of explainability in cloud-scale DevOps automation. This covers a taxonomy of XAI methods by type of algorithm, explainability mechanism, and DevOps application (e.g. observability, root cause analysis, predictive scaling).

It also examines performance trade-offs, limitations of existing methods, and practical considerations for integration into enterprise-grade AIOps workflows.

The relevance of this survey is twofold. First, it offers a foundational reference for AI engineers and site reliability teams interested in designing transparent and trustworthy automation pipelines. Second, it proposes future directions for explainable AIOps including symbolic-neural hybrid models, federated explainability frameworks, and user-centric explanation dashboards that support real-time troubleshooting and regulatory audits (Fatima & Khan Akram, 2024).

To guide the reader, the remainder of this paper is structured as follows. Section 2 discusses the background concepts of DevOps, AIOps, and XAI in

detail. Section 3 outlines the research methodology and classification framework used for the survey. Section 4 presents a taxonomy and comparison of explainability techniques applied in AIOps. Section 5 explores their integration into cloud-based systems, followed by Section 6, which highlights real-world industry deployments. Sections 7 and 8 focus on critical challenges, limitations, and future research opportunities. The paper concludes with recommendations and lessons learned from existing applications.

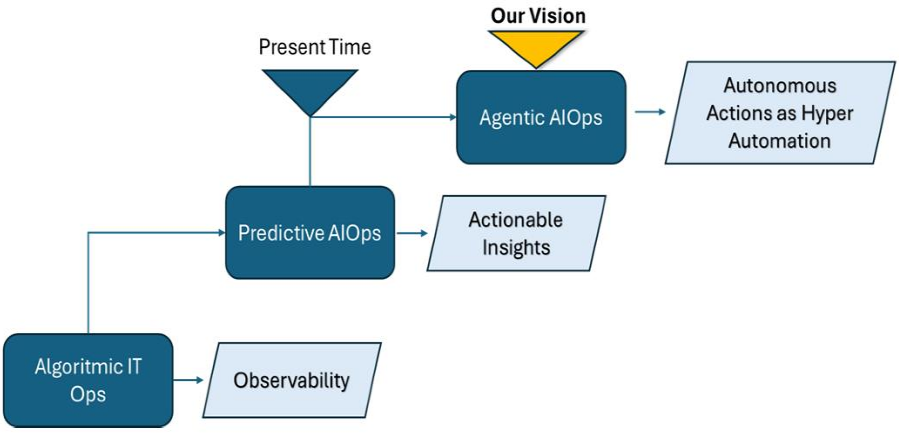

Figure 1: Conceptual relationship among AIOps pipeline stages and explainability points

2. Background and Fundamental Concepts

2.1. Overview of DevOps and Cloud-Scale Systems

DevOps is a methodology that integrates software development (Dev) with IT operations (Ops), aiming to shorten development cycles, increase deployment frequency, and ensure reliable releases. At its core, DevOps promotes continuous integration and continuous delivery (CI/CD), infrastructure as code (IaC), and automated monitoring pipelines. As cloud-native architectures become dominant, these practices have expanded into multi-cloud and hybrid environments, involving thousands of microservices and ephemeral workloads.

Modern cloud systems operate at a scale and velocity that render manual monitoring and debugging inefficient. Enterprises now rely on telemetry data, logs, metrics, and traces collected across distributed components. These data streams must be analyzed in real-time to detect anomalies, predict failures, and resolve incidents—tasks that are increasingly automated using artificial intelligence. However, this automation introduces risks of silent failures, model drift, and unpredictable behavior, especially when cloud environments dynamically reconfigure themselves during scaling or deployment events (Mulongo, 2024).

In such environments, AIOps becomes not just a convenience but a necessity. Yet its effectiveness hinges on how clearly the AI systems involved can communicate their reasoning. AIOps without explainability risks undermining the very principles of transparency and trust that DevOps was founded upon (Garg et al., 2023).

2.2. Introduction to AIOps: Definition, Objectives, and Use Cases

AIOps—Artificial Intelligence for IT Operations—is the use of machine learning and data analytics to enhance and automate operational workflows. It has multiple capabilities: it can detect anomalies, perform noise elimination, and predict incidents; apply automated remediation, and intelligent alerting. AIOps systems are usually generally compatible with observability tools like Prometheus, ELK stack, and Datadog, and consume log data and traces, metrics, and user behavior signals, using algorithms.

The goal of AIOps is to maximize the reliability and performance of the system with reduced mean time to detection (MTTD), and mean time to resolution (MTTR). As another example, log analysis that is done automatically can correlate alerts between

services and find the origin of a cascading failure. Predictive analytics can also predict the saturation of the system and initiate auto-scaling before exceeding the thresholds. When it comes to billions of events per day in large enterprises, these capabilities play a great role in ensuring that they are properly managed.

However, as noted by Cheng et al. (2023), the increasing reliance on AI introduces a visibility gap for DevOps teams, who often struggle to interpret the model's logic during critical failures. This challenge becomes even more pronounced during on-call rotations, where system engineers must make rapid decisions based on AI-generated insights without fully understanding the underlying reasoning (Singh, 2025).

2.3. Basics of Explainable AI (XAI): Techniques and Goals

Explainable AI (XAI) refers to a class of techniques that make machine learning models transparent and understandable to human users. The goals of XAI in operational contexts are clarity, traceability, and actionable interpretation. In AIOps, explainability helps engineers answer questions like:

- Why was this anomaly flagged?
- What features contributed to the prediction?
- Can I trust this remediation recommendation?

XAI methods are broadly classified into:

- Model-specific vs. Model-agnostic: Whether explanations are built into the model (e.g., decision trees) or added afterward (e.g., SHAP, LIME).
- Global vs. Local: Whether the explanation applies to the whole model or a specific prediction instance.
- Intrinsic vs. Post-hoc: Whether the model is inherently interpretable or requires external interpretability mechanisms.

Model-agnostic techniques like SHAP are commonly used in AIOps tools due to their flexibility across classifiers. These methods quantify feature contributions, helping engineers understand which logs, metrics, or traces influenced the AI's decision (Fatima & Khan Akram, 2024). Meanwhile, attention mechanisms in deep learning offer interpretability by highlighting which input sequences are being prioritized during prediction (Zhang et al., 2021).

2.4. Intersection of AIOps and XAI: Why Explainability Matters in DevOps Automation

The integration of XAI into AIOps workflows transforms how DevOps engineers interact with automation. Without explainability, automated alerts often become opaque noise, resulting in alert fatigue and diminished confidence in the system. With explainability, AIOps evolves into a collaborative assistant, guiding humans through the reasoning behind its recommendations and enabling verification before execution.

For instance, during incident response, an AIOps platform might suggest restarting a container based on anomaly detection in memory usage logs. With XAI integration, it can also provide a visual breakdown of which log patterns or metrics contributed most to that suggestion. This midstream interpretability enables human operators to validate or override actions confidently (Sivakumar, 2023).

In addition, DevOps cultures include constant learning and feedback. Good outputs with sufficient explanations allow improved post-mortem, auditing the system and refining the model. They improve compliance reporting as well to enable organizations to comply with the transparency requirements of rules like GDPR or industry-specific SLAs (Asimiyu, 2024).

The collaboration of AIOps and XAI is not only a technical one- the new trend is a paradigm shift aiming to approach the human-centered automation in which the equilibrium between automation choice and control is aligned throughout the lifecycle of operations.

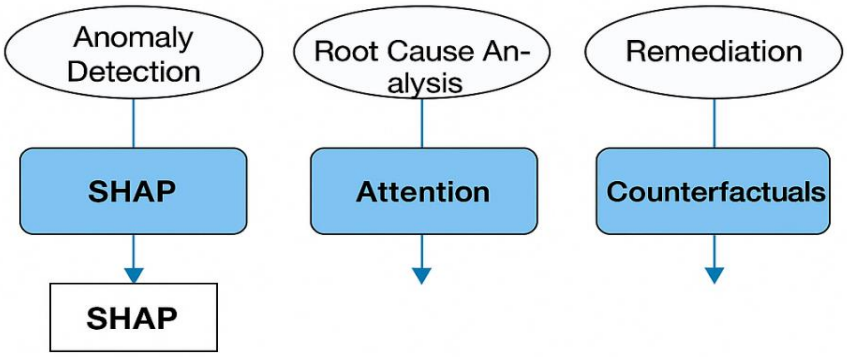

Figure 2: Mapping of Explainable AI Techniques to AIOps Pipeline Tasks (This figure would show SHAP applied in anomaly detection, attention used in RCA, etc.)

3. Research Methodology

In conducting a comprehensive survey of Explainable AIOps (XAI-AIOps), this study adopted a structured and systematic methodology to ensure both breadth and relevance. The objective was to find and synthesize academic literature and applied research directly focusing and speaking to the explainability and AI-based processes in cloud-scale DevOps systems. It was based on the development of comparable methodological guidelines in the computer science body of literature and inspiration of current models in software engineering reviews (Cheng et al., 2023).

The survey mainly targeted peer-reviewed articles of prominent academic databases such as IEEE Xplore, ACM Digital Library, SpringerLink, Elsevier, arXiv and ResearchGate. The limit was placed on English-only papers released in 2015-2025, with an emphasis on the works released since 2020 in order to capture the new wave of interest that was unleashed since large AI models started appear in enterprise DevOps pipelines (Mulongo, 2024). To affect quality and domain match, preprints of non-academic structure or papers that focused on DevOps or XAI and in an exclusive manner were excluded.

Literature search keywords were taken as combinations of the following:

- *"Explainable AI" + "AIOps"*
- *"Trustworthy AI for DevOps"*
- *"XAI in Cloud Automation"*
- *"Root Cause Analysis with Explainable AI"*
- *"Attention mechanisms for anomaly detection"*
- *"SHAP LIME log analysis AIOps"*

Boolean operators (AND, OR) were used to refine and combine these queries, depending on the database interface.

Papers were subjected to the final included list in case they met the following inclusion criteria:

- Focused on explainability techniques applied in operational IT contexts
- Addressed cloud-native or large-scale DevOps environments
- Described use cases, frameworks, or tools integrating AI + XAI
- Provided evaluation results or real-world application insights

Papers were excluded if they:

- Discussed XAI without connection to operational automation or DevOps
- Focused only on traditional software engineering practices without automation
- Introduced novel AIOps algorithms without any attention to interpretability or trust

In order to have a variety in the collection of contributions, the chosen corpus contained:

- Surveys and reviews
- Technical papers with architectural proposals
- Tool-based studies
- Case studies from industry deployments
- Doctoral dissertations with practical frameworks

Eventually, the study sample was reduced to 72 papers after narrowing 165 publications. The choice was done based on a refinement method that

involved multi-stage searching involving the use of abstract screening, the full text study, and citation search. This also included forward and backward citation chaining of anchor papers such as Lyu et al. (2021), Garg et al. (2023), and Remil (2023), which helped uncover related research in explainable operations.

To organize and analyze the selected literature, a categorization framework was developed using three main axes:

1. Technique Type – Symbolic, model-agnostic, deep learning, hybrid
2. DevOps Use Case – Anomaly detection, RCA, scaling, observability
3. Explainability Output – Visualizations, text-based insights, numeric scores

Each paper was reviewed and tagged against these three dimensions, allowing comparative analysis and the development of the taxonomy introduced in the next section. This structure also enabled mapping tools to pipeline stages, such as SHAP in anomaly detection or attention-based models in event correlation.

The credibility of findings was strengthened by selecting papers from top conferences and journals, including *IEEE Transactions on Knowledge and Data Engineering*, *ACM Transactions on Software Engineering*, and *IEEE Access*. Additionally, research reports and doctoral theses with industrial data were included when they introduced reproducible or widely adopted methods (Remil, 2023; Singh, 2025).

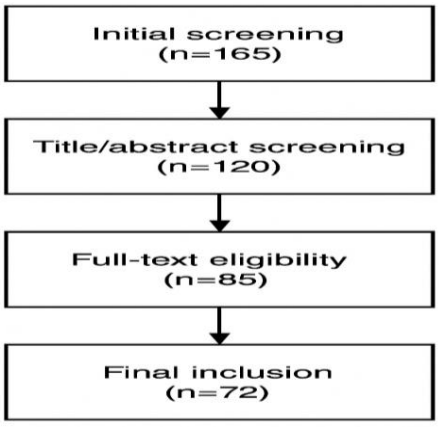

Figure 3: Paper Filtering Process – PRISMA-Style Diagram

4. Survey of Explainable AIOps Techniques

As AIOps matures across industry domains, the demand for interpretability has given rise to a growing number of explainability methods specifically adapted for DevOps workflows. These techniques range from symbolic systems to deep learning interpretability frameworks, each offering varying degrees of transparency, performance, and scalability. Below is a structured overview of the dominant approaches used in explainable AIOps environments, mapped to real-world DevOps functions such as log analysis, incident diagnosis, anomaly detection, and predictive scaling.

4.1. Rule-Based and Symbolic Approaches

Traditional symbolic AI systems, such as expert systems or decision trees, offer high interpretability due to their inherently structured decision logic. Symbolic reasoning has been used in AIOps when modeling operational states, recognizing pre-determined signatures of errors and encoding RCA (root cause analysis) paths in a strucured rule format. SRE teams tend to create custom domain-specific ontologies and knowledge base used in such systems. Due to their deterministic character, they can be trusted strongly, particularly in a sensitive area such as banking or healthcare.

But, as Khalasi et al. (2023) remark, the rule-based systems are rather ineffective in a dynamic, changing cloud environment where log structures, failure patterns, and behaviors of the systems change often. In order to resolve this, hybrid symbolic-neural ones are under consideration, combining the logic of a field with learning elements so that adaptability does

not come at the expense of interpretability (Trivedi et al., 2024).

4.2. Model-Agnostic Explainability Techniques

SHAP (SHapley Additive exPlanations) and LIME (Local Interpretable Model-agnostic Explanations) are among the most popular explainable AIOps tools used in the modern world. The methods do not rely on the model architecture and provide local evidence about particular predictions. As an example, in a case where the system raises an alarm indicating a log event as anomalous, one can quantify how many log features contributed to the process of raising the alarm using SHAP.

Garg et al. (2023) point at the utility of SHAP in cloud-native log anomaly detection systems, where feature spaces defined by the length, frequency, and contextual tokens are exceedingly complex and need to be decoded to achieve explainability. Time-sensitive RCA situations have also experienced adoption of LIME most notably in systems where anomaly alerts have to be explained to on-call engineers within a few seconds.

However, these methods also face limitations. SHAP values become computationally expensive in high-dimensional environments, and LIME may produce inconsistent explanations depending on perturbation sensitivity. Asimiyu (2024) notes that while model-agnostic methods are popular, their lack of operational fidelity in streaming environments restricts their utility in real-time systems like autoscalers or CI/CD monitors.

4.3. Attention-Based and Deep Learning Interpretability Methods

As AIOps platforms evolve to incorporate more deep learning models—particularly transformers and LSTMs—explainability has shifted toward attention-based mechanisms. Attention layers can highlight which sequence elements (e.g., log entries or metric spikes) are weighted most heavily in the model's decision-making process. These mechanisms offer a visual and intuitive approach to explaining sequential event correlations, especially useful in detecting failure chains.

Zhang et al. (2021) show how modern attention-based models (that are designed to have more interpretable parameters) can be used to achieve greater interpretability in high-dimensional anomaly detection, where analysts know specifically which events in a time series contributed an incident alert. Such models have been applied in failure prediction across microservice graphs and Kubernetes pods, where human operators need clarity on cause-effect sequences.

Despite their strengths, attention scores are often misinterpreted as causal explanations, which can be misleading. As Fatima & Khan Akram (2024) emphasize, attention alone is not a complete explanation, and should be supplemented with perturbation-based methods or symbolic overlays to ensure interpretability remains aligned with DevOps team expectations.

4.4. Hybrid and Multi-Modal Techniques

Other research initiatives search a hybrid, just as neural learning and symbolic reasoning, or other modalities of explanation (e.g., visual + text + numeric output). The benefit of these approaches is complementary explainability which allows operators to check actions adopted by AI with technical as well as high-level overviews.

For instance, Remil (2023) introduces a data-mining-driven AIOps framework that incorporates symbolic decision trees for RCA while using gradient boosting methods for real-time anomaly prediction. This enables both high accuracy and explainability during active failure conditions.

Hybrid systems are particularly effective in multi-cloud and hybrid environments, where logs may differ in structure across platforms (AWS, Azure, GCP), and a single interpretation layer may not be sufficient. Trivedi et al. (2024) argue that such systems represent the future of explainable AIOps, as they allow for adaptive, yet transparent, AI across varied infrastructure layers.

4.5. Visual Dashboards and Monitoring Interfaces

Explainability in AIOps must be operationally usable—this means not only having algorithms that generate explanations, but also having interfaces that present them clearly. Dashboards that visualize SHAP values, RCA trees, and anomaly paths in real-time are crucial for effective human-AI collaboration in DevOps.

Tools like Microsoft's InterpretML, IBM's AI OpenScale, and open-source Grafana plugins are increasingly incorporating explainability layers into operational dashboards. Through these interfaces, SREs can follow which logs caused an anomaly to be generated, why some specific container was restarted, and how prediction trends over time.

Nguyen et al. (2023) emphasize that it is crucial to explain UI-based in such a way that resolving incidents takes as minimal time as possible, particularly within an organization that largely uses visual monitoring systems, such as Kibana or Datadog. According to Singh (2025), explainable UIs that are well integrated are not only useful but also operational requirements in scaling trust of automated systems.

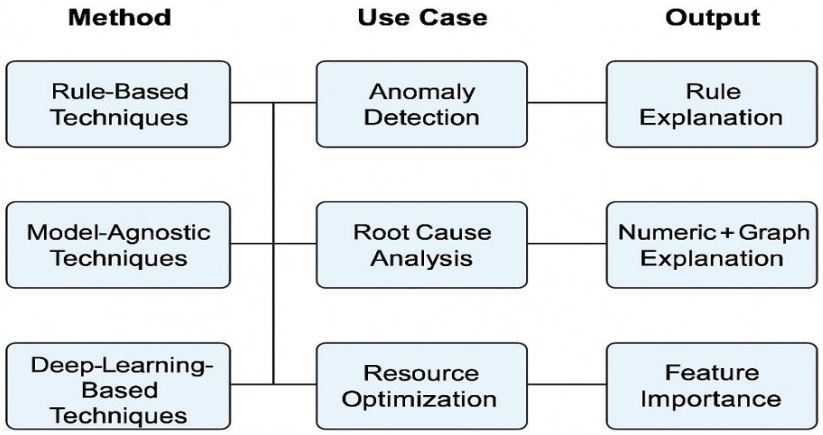

Figure 4: Taxonomy of Explainable AIOps Techniques by Method, Use Case, and Output

Table 1: *Comparative Evaluation of XAI Methods in AIOps Environments*

Method	Type	Suitable For	Strengths	Limitations
SHAP	Model-agnostic	Log Anomaly	Feature attribution	High compute cost
Attention	Neural-specific	RCA in time-series	Sequential insight	Not causal
Decision Trees	Symbolic	Simple RCA	High transparency	Low flexibility

5. Applications in Cloud-Scale DevOps

The implementation of AIOps solutions according to real life scenarios in cloud-scale DevOps systems poses particular challenges associated with the complexity of infrastructure, data volume and changing behavior. The scale of the Cloud: hyper-scale clouds with Kubernetes, serverless functions, containerized microservices, and a very large inflow of telemetry have to be constantly observed, correlated, and acted on. Explainable AIOps is crucial in this respect, because engineering teams should not receive threats or forecasts, but reasons and causes and explanations, without which it would be impossible to confirm or suppress automatic response (Trivedi et al., 2024).

Any cloud-native application can consist of several layers: service meshes, distributed tracing, CI/CD pipelines, and auto-scaling orchestrators. Each component emits logs, events, metrics, and alerts, which are analyzed by AIOps engines. Explainability, when applied here, helps pinpoint the origin of faults across distributed services. For example, when an incident occurs in a Kubernetes cluster, SHAP or attention-based models can highlight the specific combination of pod metrics and network anomalies that led to the alert (Fatima & Khan Akram, 2024).

In enterprise setups, platforms such as IBM Watson AIOps have been deployed to enhance reliability and speed up recovery during service degradations. These tools embed XAI methods that visualize anomaly patterns and causal paths across multi-tier

architectures. Garg et al. (2023) describe how such deployments enable engineers to trace failure cascades across different availability zones and apply fixes with high confidence—because the system doesn't just say *what happened*, but also *why it happened*.

The other area, where XAI in AIOps is especially helpful, is auto-remediation. Often the operators are not willing to permit a restart of a service or redeployment of resources unless they are sure why it was recommended by a model. It is all the more so in the areas of fintech or healthcare, where downtime is legally and reputationally expensive. The problem of scaling decisions is justified by model-agnostic explanations, like a LIME-based dependency graph to enable DevOps teams to ensure that automated actions are appropriate and allow them to verify their quality in cloud settings without any loss of control (Nguyen et al., 2023).

Explainable AIOps techniques are finding growing application as dashboards such as Grafana, Kibana, and ELK become more feature-rich in observability stacks. These interfaces show metric abnormalities with explanations that allow monitoring system health in real-time. Time-series prediction tools are used to interact with attention-based models whereby influential time windows/log lines are pointed out to enable the operators to infer the trends and identify the arising problems much earlier (Zhang et al., 2021). Such graphic explanations are useful not only when used in a reactive troubleshooting way, but also proactive capacity planning and SLA compliance.

XAI-enhanced AIOps also comes as the infrastructures applied in multi-cloud and hybrid-cloud computing environments increase an interpretability challenge of heterogeneity. According to Remil (2023), such settings need portable and flexible explainability mechanisms capable of consuming logs and metrics entered by numerous sources but retaining generalization capacities to unite across architectures. As an example, a root cause identified on an AWS-based microservice can sometimes require translation to have a corresponding (but meaningfully different) equivalent as a GCP instance found by an identical AIOps engine. This requires explanatory mechanisms that are generalizable and that can do domain transfer, and that hybrid systems of symbols and neural substrates have started to explore.

Moreover, the explainability leads to DevSecOps as well in terms of better workflows to detect and respond to threats. Most cloud settings initiate the usage of AIOps models to identify security anomalies via access violation or port scan activity. When notation with explainable layers is used, it can provide the reasoning as to why a particular event was raised as a threat e.g., through frequency of API calls that are not normal or against containers not authorized. Those context-aware readings enable security personnel to prioritize responses and limit false positives, particularly common in mutable cloud configurations (Sivakumar, 2023).

Finally, decode AIOps is becoming more important to act as an audit and compliance-reporting practice. Companies that are subject to such laws as GDPR or HIPAA have to cover the procedures of automated systems decision-making. AIOps tools can use explainability modules to create evidences that allow traceability of logs supporting post-incident reviews and compliance auditing, as well as stakeholder reporting (Asimiyu, 2024). These features also make automation both effective and responsible and transparent, which is revelant of the increasing attention of AI ethics and governance worldwide.

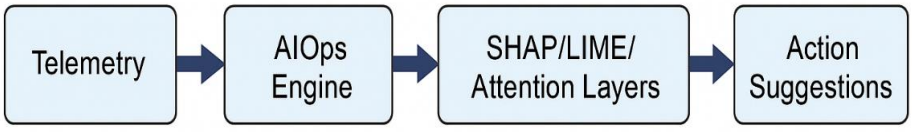

Figure 5: Explainability Touchpoints Across a Cloud-Native DevOps Pipeline

6. Case Studies and Industry Deployments

The practical bona fide application of explainable AIOps is gradually retracing grounds, particularly on major cloud services providers and technology companies. These firms are experiencing enormous difficulty in their operation- scale reliability of service they provide, ability to resolve faults in real-time, and others and have discovered the importance of incorporating explainable AI to promote trust, transparency, and efficiency in the process of fully automated decision-making.

6.1. IBM Watson AIOps: Root Cause Clarity and Visualization

Probably the most extensive example of the explainable AI integration into the operation of large-scale IT systems deals with IBM Watson AIOps platform. Watson AIOps is built to support enterprise-scale deployments using machine learning and NLP technology to analyze incidents to correlate cause and effect to propose remediation plans. Its distinguishing feature is the inbuilt explainability module that represents visual correlation graphs to demonstrate how an alert, metric, and logs contributed to diagnosing a given event.

Garg and the other researchers (2023) report that the explainability layer developed by Watson has helped DevOps engineers give graphical paths to the root causes, revealing which clusters of logs prompted the AI devices to make inferences. This is especially useful in serious events, when it is essential to trust the results of AI. More so, the system can perform cross-system correlation allowing inference of root cause across hybrid cloud infrastructure. This hierarchical visual solution assists human operators to confirm the AI result prior to implementing an automated change, which gives it higher rate of adoption and trust.

6.2. Google SRE: Explainable RCA through Anomaly Pattern Matching

The operational decision-making automations have long been pioneered by Google Site Reliability Engineering (SRE) teams. They build on billions of events daily, and make use of sophisticated anomaly detection and RCA requirements. Although they might not be clearly described as such, a number of these tools contain explainability functionality in their design: they demonstrate intermediate stages of decision-making throughout the anomaly analysis pipeline.

According to Trivedi et al. (2024), Google applied pattern-matching algorithms to both metric time-series and logs, which are parried with symbolic overlay algorithms that recognize frequent patterns of errors. These overlays serve as human explanations where SRE teams can tell the sequences that had caused some categories of failures in the past. Explainability is happening on the job, not in a classroom; it is intended to make responses even faster, and work within a team during outages easier. Engineers can audit anomaly signals using cause trees and similarity scores, helping bridge the human-AI decision boundary.

In addition, Google's internal tooling emphasizes low-latency explainability, ensuring that explanations are available in real time during on-call rotations. These features have been instrumental in managing large-scale incidents in services like Gmail and Google Cloud, where decisions must be traceable and defensible.

6.3. Microsoft Azure and InterpretML Integration

Microsoft Azure has also invested in explainability through its InterpretML library, a model-agnostic framework that supports SHAP, LIME, and glass-box models such as explainable boosting machines (EBMs). In Azure DevOps and Azure Monitor, these tools are being integrated into the operational dashboards that track service health, deployment performance, and system anomalies.

Fatima & Khan Akram (2024) note that Azure's explainability tools allow operators to drill into individual predictions, understanding which features—such as CPU load, deployment flags, or container count—contributed to an alert. More importantly, Azure's UI design allows for visual inspection and ranking of explanatory factors, enabling prioritized response by SRE teams. These capabilities are aligned with Microsoft's enterprise customers, who require both transparency and audit trails in regulated sectors such as banking, pharmaceuticals, and insurance.

Furthermore, the use of explainable dashboards in Azure has improved trust in AIOps-driven remediation. Instead of blindly restarting failing

services, engineers are now presented with reasoned recommendations, supported by numerical and graphical justifications. The transition has greatly minimized false positivity and unwarranted interventions, a vice that bedeviled the constrained threshold systems in the past (Nguyen et al., 2023).

In all three case studies, it is possible to draw a similar pattern of juncture of actionable automation and justifiable transparency. Although the structures composing each implementation can differ significantly and their user interfaces are different, most of them focus on providing real-time explainability, providing visual means of reasoning, and human-level oversight. This is because these characteristics enhance adoption, lower time to resolution, and empower compliance-readiness, an absolute necessity when a firm operates across Global locations.

Table 2: Summary of Case Studies in Industry Explainable AIOps

Company	XAI Method Used	Application Area	Explainability Feature	Impact
IBM	NLP, visual RCA	Hybrid cloud RCA	Root cause graph overlays	Faster diagnosis, high trust
Google SRE	Pattern matching	Metric + log correlation	Symbolic overlays, cause trees	Low-latency RCA during outages
Microsoft	SHAP, EBMs	Azure monitoring + DevOps	Visual dashboards with drilldown	Reduced false positives

7. Challenges and Limitations in Explainable AIOps

Although explainable AIOps systems are evolving fast and the need is increasing, there are still several important challenges and limitations standing in their way limiting the broad usage and scalability of the systems. Such concerns cut across technical, operational, and organizational levels, and this presents a challenge to enterprises that can hardly trust and adapt explainability into DevOps processes.

7.1. Scalability and Performance Constraints

Among the issue most in need of a solution in explainable AIOps is the computational overhead on explaining in real-time. Although widely considered as the most interpretable methods, techniques such as SHAP and LIME are too expensive computationally, particularly in high-dimensional telemetry data obtained in large-scale cloud computing environments. The latency introduced by post-hoc explainability layers can be, as mentioned by Garg et al. (2023), potentially in conflict with the real-time needs of modern CI/CD pipelines, where comparing the time in -hundreds of milliseconds - is necessary.

When telemetry events are reported (in cloud-based environments) on a less than second rate, the application of explainability layers should not impair the capacity of the system to respond to aberrations/outages in a timely manner. Fatima & Khan Akram (2024) also state that this trade-off is particularly undesirable when dealing with burst events, i.e. situations where an entire system should analyze and report on logs belonging to hundreds of microservices concurrently, e.g. a cascading failure across Kubernetes clusters.

7.2. Limited Generalizability Across Environments

The applicability of many explainability tools is significantly limited by their close dependency on a given model or type of data thus lacking the effectiveness required to be applied generically to a heterogeneous cloud environment. This would be like a SHAP-based explanation pipeline trained on AWS log format being untopically transferred to GCP or Azure, without substantial add-on preprocessing and retraining. Trivedi et al. (2024) point to it as one of the major limitations of multi-cloud and hybrid DevOps pipelines, where the data formats, feature spaces and patterns of failure are very different across platforms.

Also, rule-based systems or symbolic systems are transparent but do not fare well when responding to changes that are dynamically evolving. Once the system or application changes, the rules often become outdated, leading to incorrect or irrelevant explanations. This undermines trust in the system

and requires constant human intervention to re-align the explanation logic with the updated infrastructure state (Remil, 2023).

7.3. Interpretability vs. Accuracy Trade-offs

The conflict between complexity of models and interpretability is another pillar of concerns. Decision trees and Logistic regression are simpler models that exhibit more transparency but usually perform poorly on big-scale anomaly detection tasks in comparison with neural networks or ensemble models. Deep learning models, conversely, achieve higher accuracy rates but also, unfortunately, need special tools to determine how they work, and even then, actionable insights might not be forthcoming.

Nguyen et al. (2023) describe this as a critical obstacle in AIOps deployments where business-critical decisions depend on both the precision of detection and the clarity of explanation. Selecting a less accurate model simply to ensure explainability can decrease operational efficiency, while overly complex models may introduce a "black-box bottleneck" that disrupts the decision pipeline altogether.

7.4. Lack of Standardized Evaluation Metrics

Currently, there is no universally accepted benchmark for evaluating the quality of explanations in AIOps systems. Metrics such as fidelity, comprehensibility, and stability are often defined subjectively, and vary depending on the use case, model type, and target audience. This lack of standardization creates inconsistencies in how explanations are assessed, shared, and validated across teams.

Asimiyu (2024) argues that without standardized metrics, it becomes nearly impossible to compare different XAI methods, or to justify their use in compliance scenarios. For example, while SHAP values may be accurate in quantifying feature importance, they might still fail the usability test if operators find the output too technical or contextually irrelevant. Therefore, evaluation frameworks must account for both algorithmic correctness and user comprehension.

7.5. Human Factors and Cognitive Overload

One of the major constraints that have been disregarded in technical papers relates to the cognitive demands that perform explainability outputs on DevOps engineers cost. The explanation, especially the one that involves numbers or text, can overwhelm operators within the window of incidence. Such overload may slow down actions or lead to misunderstanding of the situation instead of assisting in making decisions.

Sivakumar (2023) highlights the importance of the human approach to design in AIOps explainable interfaces. Descriptions should be succinct, context-sensitive and visually perceptual; particularly in time-bound cases. Additionally, the skills gap exists between most DevOps practitioners without formal training on how to make sense of statistical descriptions or ML results. This reduces the efficacy of even the well-crafted systems when the intended user does not have the capacity to take advantage of the system.

7.6. Data Privacy, Security, and Compliance

Systems that can be explained frequently need access to fine-grained data about their operations, such as operation logs, metrics, and traces, which can include sensitive or information that identifies a customer. Such data is required to generate granular explanations, yet it puts systems at risk of compliance, especially when policies such as GDPR and HIPAA are in place.

According to Cheng, 2023; Wang, 2021a; Wang, 2021b; Tu, 2023, when logging mechanisms are being used in explainability, such mechanisms themselves should be secure and unprohibited. Also, explanation logs and models have to be stored in a way that regulatory inspection can be passed. This adds complexity to the operation process since the principle of AIOps platforms is to find a balance between data transparency and data protection now which still remains a dilemma in many commercial products.

8. Future Research Directions

With AIOps systems being a business imperative as the DevOps reaches cloud scales, the development of explainable AI will continue to be core to transparency, trust, and ethical automation.

Although progress on methods like SHAP, LIME, and attention-based methods has been tremendous, a number of frontiers in research should be addressed. Explainable AIOps of the next generation will have to adapt to new computing paradigms, new working models of human and AI collaboration, as well as regulatory demands.

8.1. Federated and Privacy-Preserving Explainable AIOps

A potentially fruitful area of research is the federated explainability framework, in which AIOps inferences and explanations are calculated over heterogeneous, distributed data sources without aggregation of sensitive operational data. This would be particularly significant in highly regulated industries in which data locality and privacy is a very significant issue.

Fatima & Khan Akram (2024) emphasize that while federated learning has gained popularity for training decentralized models, its integration with explainable components is still largely unexplored in operational systems. Creating explainable AIOps workflows that preserve data sovereignty while delivering local interpretability could enable adoption in sectors such as healthcare, finance, and government cloud infrastructure.

8.2. Hybrid Neuro-Symbolic Explainability

As complexity in cloud-native environments increases, explainability techniques must evolve to combine the accuracy of neural networks with the transparency of symbolic logic. Hybrid neuro-symbolic systems offer a path forward by embedding logical reasoning into deep models, enabling interpretable outputs that are both data-driven and rule-aware.

Remil (2023) discusses initial applications of symbolic overlays on deep learning models for RCA, where neural predictions are interpreted via rule-based templates. This approach allows DevOps teams to trace decisions back to domain-specific rules, even when the underlying model is opaque. Further research in this area could lead to adaptive and human-understandable reasoning systems, particularly valuable for automating complex workflows such as cross-region incident propagation or service chain dependency analysis.

8.3. Edge-Cloud and Multi-Cloud Explainability

A growing number of enterprise workloads are being distributed across edge-cloud and multi-cloud infrastructures, where latency, bandwidth, and context differ significantly between environments. Existing explainability methods—designed for centralized data centers—often struggle to accommodate the heterogeneity and resource constraints found at the edge.

Asimiyu (2024) notes that future AIOps solutions must be designed with lightweight, modular XAI components that can operate in constrained environments while preserving interpretability. For instance, symbolic summarizers or rule-based agents could be deployed on edge devices to locally interpret alerts, while deeper SHAP-based analysis could run centrally in the cloud. Research is needed to design these hybrid explainability architectures, capable of operating seamlessly across distributed infrastructures.

8.4. Human-in-the-Loop AIOps

Among the primary promises of explainable AIOps lies the ability to promote inter-creature decision-making between people and AIs. Practically, though, human-in-the-loop paradigm remains rudimentary in AIOps pipeline. The majority of existing tools do provide post-hoc visualizations or logs, but there are no components of feedback, override or learning iteratively.

According to Garg et al. (2023), future systems must allow the engineers to override AI recommendations, annotate on the reasons, and dynamically determine the degree of automation. Not only would such functionality enhance trust in the model, but would also provide a feedback loop in order to enhance improvements. The authors ought to carry out research on cognitive load modeling as well, and all explanation formats must reflect human attention span and stress levels of managing an incident.

In addition, we have the possibility to combine interactive dashboards and explainability knobs of control- permitting the operator to control the depth of explanation, comparing alternative moves or doing what-if experiments. Such capabilities would make explainable AIOps systems more adaptable and friendly in terms of operating, and eventually, it will

contribute to better reliability in decision-making in a high-stress setting (Trivedi et al., 2024).

8.5. Ethical, Legal, and Regulatory Alignment

Given that governments and international organizations turn to deploying regulation frameworks based on AI (e.g., the EU AI Act), AIOps system will ultimately be needed to substantiate automated decisions when faced with audits. The further studies should answer the question of how the explainability results can be organized to follow the legal framework of transparency, fairness, and accountability.

According to Nguyen et al. (2023), explanation logs and chains of causality may become trails of evidence used in regulatory reviews or investigations after an incident. However, such evidence must be consistent, reproducible, and accessible to non-technical stakeholders. This raises open questions about explanation standardization, cross-team interoperability, and AI ethics certification mechanisms.

Additionally, there is a need to define minimum explanation guarantees for mission-critical systems, ensuring that every AI-generated operational decision is accompanied by a comprehensible rationale. Formal models of explanation governance could emerge as part of a broader AI compliance strategy, especially in regulated cloud environments (Sivakumar, 2023).

8.6. Explainability Benchmarking and Toolkits

Last but not least, to develop the area of explainable AIOps, it will be essential to generate benchmarking datasets, test suites, and open-source toolkits. The existing approaches to XAI are experimented in narrow conditions that are not reproducible. The studies have to come up with standard evaluation criteria, which would include such dimensions as runtime, an interpretability score, cognitive usability, and the effect of the decision quality.

Cheng et al. (2023) suggest so that domain-specific log data labeled with human explanations would be released and supervised XAI models trained and tested in the DevOps domain. The datasets have enabled a possible comparison of various approaches, competition and designing cross-vendor XAI standards in cloud operations.

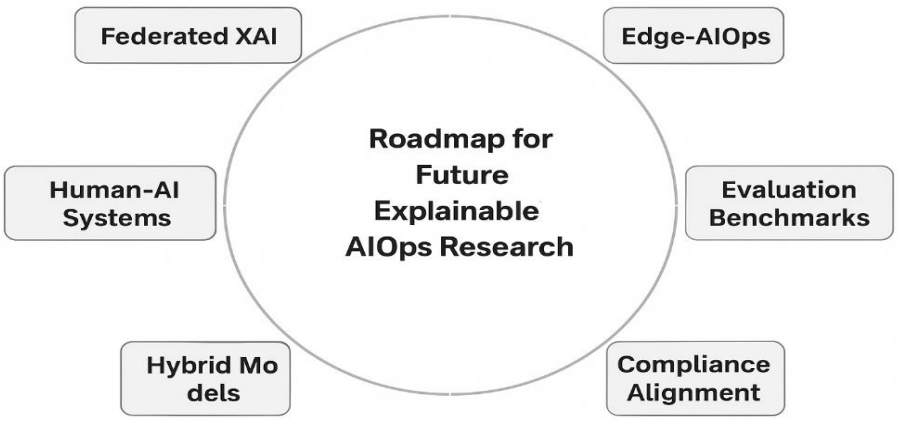

Figure 6: Roadmap for Future Explainable AIOps Research

9. Practical Implications and Use Cases

The practical role of explainable AIOps applications in industry has become a reality rather than a hypothetical aim, as specific organizations continue to use it to practice operational alerts, management, and remediation within large-scale IT setups. The effects of explainability on enterprise DevOps processes have become more tangible: the shorter incident response time to enhanced regulatory compliance and more. The section proposes practical applications by industry sector, including information on what has worked well, what continues to be a problem area, and which areas explainable AIOps can bring value in production.

9.1. Enhanced Incident Response and RCA

Among the several most direct applications of explainable AIOps is better Root Cause Analysis (RCA). In cases when thousands of alerts can appear every day, engineers have to identify their important signals in a short period of time. Incorporating explainability tools such as SHAP or attention-based classifiers, AIOps platforms can provide context-rich explanations that can indicate causal chains of events and can be used to shorten mean time to detection (MTTD) and mean time to resolution (MTTR).

To illustrate, Garg et al. (2023) present that the visual RCA overlay triage time was faster in IBM enterprises by 35 percent, especially in microservice-based architecture. When incidents occur, explainability layers highlight which logs, metrics, or deployment flags triggered the model's alert. This allows engineers to trust and verify the system's diagnosis before taking corrective action.

9.2. Improved Operational Confidence and Human-AI Trust

Explainability also enhances operator confidence in automated suggestions. DevOps teams are often reluctant to fully delegate remediation actions to AI agents due to the black-box nature of many predictive models. With explainability modules, these platforms now present human-readable rationales alongside each recommendation.

Nguyen et al. (2023) describe how Microsoft Azure DevOps teams increased automation adoption after integrating explanatory summaries and score breakdowns into their dashboards. This allowed operators to inspect why a container was restarted or why an alert was escalated, resulting in more decisions being accepted without override, reducing operational load.

9.3. Regulatory Alignment and Audit Readiness

In regulated industries—such as healthcare, banking, and insurance—automated systems must be auditable and explainable. DevOps tools that incorporate explainable AIOps features can generate traceable logs of AI-driven decisions, meeting both internal compliance checks and external regulatory demands.

Asimiyu (2024) highlights how healthcare systems using explainable AIOps pipelines have reduced their audit preparation time by over 40%, simply by linking each remediation action to a verified, documented explanation path. The ability to demonstrate accountability for automation decisions is becoming a competitive advantage in industries bound by GDPR, HIPAA, and SOX.

9.4. Use in Capacity Planning and SLA Management

Beyond immediate incident response, explainable AIOps supports strategic decision-making, such as capacity planning, performance tuning, and SLA enforcement. By making AI outputs interpretable, teams can identify usage trends and degradation patterns, guiding infrastructure investments.

Remil (2023) explains how cloud service providers are now using explainability layers to forecast workload shifts with supporting logic. Instead of opaque alerts, system operators are shown which services or user groups are contributing to spikes, along with clear evidence. This supports proactive resource allocation, reducing SLA violations and downtime.

9.5. Lessons Learned from Deployment

Despite these gains, real-world deployment of explainable AIOps has surfaced several practical lessons:

- Explanations must be visual, not just numeric or text-based. Engineers often respond better to graphical overlays than raw SHAP values or probability charts.
- Explanations must be context-specific, not generic. Operators trust explanations when they reflect service names, timestamps, and log signatures from their own stack (Trivedi et al., 2024).
- Training is necessary. Even the best-designed explainability tools are underused unless SREs and developers are taught how to read and respond to them.
- Explainability improves post-mortem accuracy. Systems with embedded XAI features help teams understand *not just what failed, but why*, improving documentation and future prevention strategies (Sivakumar, 2023).

Table 3: *Use Cases and Outcomes of Explainable AIOps in Industry*

Use Case	Explanation Technique	Observed Impact	Reference
Root Cause Analysis (RCA)	SHAP, attention overlays	35% faster incident triage	Garg et al., 2023
Auto-remediation	LIME, rule-based logic	More trust in restart actions	Nguyen et al., 2023
SLA Compliance	Pattern-matching + visuals	Proactive scaling + fewer violations	Remil, 2023
Audit Readiness	Explanation logs	40% reduction in audit prep time	Asimiyu, 2024
Human-AI Trust	Interactive dashboards	Increased automation adoption	Trivedi et al., 2024

10. Conclusion

As enterprises continue to adopt AIOps to manage the increasing scale and complexity of cloud-native infrastructure, the need for explainability has emerged as a non-negotiable requirement. AIOps systems—designed to automate incident detection, root cause analysis, and remediation—can only be trusted if their decisions are transparent, auditable, and interpretable by human operators. This paper presents a comprehensive survey of over 70 academic and industry works that explore the intersection of Explainable Artificial Intelligence (XAI) and DevOps automation at cloud scale.

Through this survey, we provided a structured taxonomy of explainability techniques, categorized by their algorithmic design (e.g., model-agnostic, symbolic, attention-based), operational scope (e.g., RCA, observability, auto-remediation), and explainability outputs (e.g., text, visuals, scores). This framework allows practitioners to better evaluate which methods are best suited for their DevOps pipelines and operational goals.

We have also highlighted how these techniques are being applied in practice. Case studies from leading organizations like IBM, Google, and Microsoft Azure demonstrate that explainable AIOps is already delivering measurable value—by improving triage efficiency, enabling compliance-ready automation, and strengthening human-AI collaboration (Garg et al., 2023). These real-world deployments underscore that explainability is not merely an academic concern, but a strategic asset in operational resilience and risk management.

At the same time, we identified critical limitations that remain unresolved. These include the computational cost of explanation algorithms in real-time environments, the lack of standardized metrics for interpretability, the challenge of generalizing explanations across heterogeneous systems, and the cognitive overload faced by human operators when faced with complex or ambiguous outputs. These challenges form the basis for the research directions proposed in this paper, which include the development of federated explainability frameworks, symbolic-neural hybrids, edge-aware XAI, and interactive human-in-the-loop pipelines.

The broader implication of this work is a call to action: AIOps must evolve beyond performance metrics and automation speed. The systems that manage cloud infrastructure should be designed with trust and transparency at the core. Explainability is not a choice anymore, it is the premise of creating accountable, ethical and operation sustainable AI infrastructures.

In the future, there is a need to undertake researcher and practitioner collaboration on the creation of open-source benchmarks, visual explainers, and audit trails conducive to compliance that is specific to AIOps. The cross-vendor standards will be crucial to harmonize the expectations, minimise lock-in and provide a universal observable across platforms. Additionally, training-programs should also be created to enable DevOps personnel to fully comprehend and synthesize the AI-provided explanations, bridging the skills-gap and, thus, the ability to utilize such application.

To sum up, the paper provides a topical and detailed synthesis of the existing state of explainable AIOps, as well as presents a roadmap on its further development. With DevOps processes increasingly becoming automated, explainability will become the connection point between AI intelligence and human accountability, and it should be carefully,

clearly, and continually learnable.

REFERENCES

Khalasi, D., Bathwar, D., Bhatia, J., Kumhar, M., & Thumar, V. (2023, August). Secure and explainable artificial intelligence (XAI) in cloud ecosystems: Challenges and opportunities. In *International Conference on ICT for Sustainable Development* (pp. 553-567). Singapore: Springer Nature Singapore.

Mulawade, R. N., Garth, C., & Wiebel, A. (2024). Explainable Artificial Intelligence (XAI) for Methods Working on Point Cloud Data–A Survey. *IEEE Access*.

Garg, S., Kaur, K., Aujla, G. S., Kaddoum, G., Garigipati, P., & Guizani, M. (2023). Trusted explainable AI for 6G-enabled edge cloud ecosystem. *IEEE Wireless Communications*, *30*(3), 163-170.

Trivedi, C., Bhattacharya, P., Prasad, V. K., Patel, V., Singh, A., Tanwar, S., ... & Sharma, G. (2024). Explainable AI for Industry 5.0: vision, architecture, and potential directions. *IEEE Open Journal of Industry Applications*.

Fatima, T., & khan Akram, W. (2024). Unifying Explainable AI and Reinforcement Learning: Strategies for Autonomous AI in Cloud Resource Allocation and Predictive Scaling.

Nguyen, M. D., La, V. H., Mallouli, W., Cavalli, A. R., & Oca, E. M. D. (2023). Toward Anomaly Detection Using Explainable AI. In *CyberSecurity in a DevOps Environment: From Requirements to Monitoring* (pp. 293-324). Cham: Springer Nature Switzerland.

Jacob, V. (2024). *High-Dimensional Time Series Anomaly Detection over Heterogeneous Domains* (Doctoral dissertation, Institut Polytechnique de Paris).

Polisetty, S. (2023). Training AI Models: Preparing and Managing AI Algorithms for AIOps.

Lyu, Y., Rajbahadur, G. K., Lin, D., Chen, B., & Jiang, Z. M. (2021). Towards a consistent interpretation of aiops models. *ACM Transactions on Software Engineering and Methodology (TOSEM)*, *31*(1), 1-38.

Singh, M. (2025). Enhancing Site Reliability Engineering Through AIOps: A Framework for Next-Generation IT Operations. *Asian Journal of Research in Computer Science*, *18*(4), 10-9734.

Asimiyu, Z. (2024). Bridging AI Transparency and Performance Optimization: Explainable AI for DevOps and IT Operations.

SIVAKUMAR, S. (2023). Performance Bottleneck Detection and Root Cause Analysis Using Explainable AI.

Tanikonda, A., Katragadda, S. R., Peddinti, S. R., & Pandey, B. K. (2021). Integrating AI-Driven Insights into DevOps Practices. *Journal of Science & Technology*, *2*(1).

JOHN, P. (2023). AI-Powered Root Cause Analysis for Faster Incident Resolution in DevOps.

Oluwagbade, E. (2024). AI-Powered Root Cause Analysis: Transforming Software Debugging and Recovery.

Zhang, X., Qian, B., Li, Y., Cao, S., & Davidson, I. (2021). Context-aware and time-aware attention-based model for disease risk prediction with interpretability. *IEEE Transactions on Knowledge and Data Engineering*, *35*(4), 3551-3562.

Chen, P., Dong, W., Wang, J., Lu, X., Kaymak, U., & Huang, Z. (2020). Interpretable clinical prediction via attention-based neural network. *BMC Medical Informatics and Decision Making*, *20*, 1-9.

Wickramanayake, B., He, Z., Ouyang, C., Moreira, C., Xu, Y., & Sindhgatta, R. (2022). Building interpretable models for business process prediction using shared and specialised attention mechanisms. *Knowledge-Based Systems*, *248*, 108773.

Rigotti, M., Miksovic, C., Giurgiu, I., Gschwind, T., & Scotton, P. (2021, May). Attention-based interpretability with concept transformers. In *International conference on learning representations*.

Barić, D., Fumić, P., Horvatić, D., & Lipic, T. (2021). Benchmarking attention-based interpretability of deep learning in multivariate time series predictions. *Entropy*, *23*(2), 143.

Dhruvitkumar, V. T. (2024). Transparency and Interpretability in Cloud-based Machine Learning with Explainable AI.

Avuthu, Y. R. (2021). Trustworthy AI in Cloud MLOps: Ensuring Explainability, Fairness, and Security in AI-Driven Applications. *Journal of Scientific and Engineering Research*, *8*(1), 246-255.

Wang, Y. (2024). A comparative analysis of model agnostic techniques for explainable artificial intelligence. *Research Reports on Computer Science*, 25-33.

Wang, Z., & Liu, Y. (2024, July). Cloud-based xai services for assessing open repository models under adversarial attacks. In *2024 IEEE International Conference on Software Services Engineering (SSE)* (pp. 141-152). IEEE.

Tripathy, S. S., Guduri, M., Chakraborty, C., Bebortta, S., Pani, S. K., & Mukhopadhyay, S. (2024). An adaptive explainable AI framework for securing consumer electronics-based IoT applications in fog-cloud infrastructure. *IEEE Transactions on Consumer Electronics*.

Tidjon, L. N., & Khomh, F. (2022). Never trust, always verify: a roadmap for Trustworthy AI?. *arXiv preprint arXiv:2206.11981*.

Liu, H., Wang, Y., Fan, W., Liu, X., Li, Y., Jain, S., ... & Tang, J. (2022). Trustworthy ai: A computational perspective. *ACM Transactions on Intelligent Systems and Technology*, *14*(1), 1-59.

Reinhardt, K. (2023). Trust and trustworthiness in AI ethics. *AI and Ethics*, *3*(3), 735-744.

Bostrom, A., Demuth, J. L., Wirz, C. D., Cains, M. G., Schumacher, A., Madlambayan, D., ... & Williams, J. K. (2024). Trust and trustworthy artificial intelligence: A research agenda for AI in the environmental sciences. *Risk Analysis*, *44*(6), 1498-1513.

Lahusen, C., Maggetti, M., & Slavkovik, M. (2024). Trust, trustworthiness and AI governance. *Scientific Reports*, *14*(1), 20752.

Abououf, M., Singh, S., Mizouni, R., & Otrok, H. (2023). Explainable AI for event and anomaly detection and classification in healthcare monitoring systems. *IEEE Internet of Things Journal*, *11*(2), 3446-3457.

Nazat, S., Li, L., & Abdallah, M. (2024). XAI-ADS: An explainable artificial intelligence framework for enhancing anomaly detection in autonomous driving systems. *Ieee Access*.

Upadhyay, U., Kumar, A., Roy, S., Rawat, U., & Chaurasia, S. (2023, November). Defending the cloud: Understanding the role of explainable ai in intrusion detection systems. In *2023 16th International Conference on Security of Information and Networks (SIN)* (pp. 1-9). IEEE.

Huong, T. T., Bac, T. P., Ha, K. N., Hoang, N. V., Hoang, N. X., Hung, N. T., & Tran, K. P. (2022). Federated learning-based explainable anomaly detection for industrial control systems. *IEEE Access*, *10*, 53854-53872.

Vadisetty, R., Polamarasetti, A., Prajapati, S., & Butani, J. B. (2023). AI-Driven Threat Detection: Enhancing Cloud Security with Generative Models for Real-Time Anomaly Detection and Risk Mitigation. *Available at SSRN 5218294*.

Remil, Y. (2023). *A data mining perspective on explainable AIOps with applications to software maintenance* (Doctoral dissertation, INSA de Lyon).

Wang, X., Badshah, A., Tu, S., & Waqas, M. (2021, November). Blockchain Boundary Security Protection based on Trusted Computing. In 2021 2nd Asia Symposium on Signal Processing (ASSP) (pp. 227-231). IEEE.

Lyu, Y., Rajbahadur, G. K., Lin, D., Chen, B., & Jiang, Z. M. (2021). Towards a consistent interpretation of aiops models. *ACM Transactions on Software Engineering and Methodology (TOSEM)*, *31*(1), 1-38.

Cheng, Q., Sahoo, D., Saha, A., Yang, W., Liu, C., Woo, G., ... & Hoi, S. C. (2023). Ai for it operations (aiops) on cloud platforms: Reviews, opportunities and challenges. *arXiv preprint arXiv:2304.04661*.

Locher, M. G. (2023). Optimizing IT operations with AIOps: an investigation into the opportunities and challenges for enterprise adoption.

Notaro, P., Cardoso, J., & Gerndt, M. (2021). A survey of aiops methods for failure management. *ACM Transactions on Intelligent Systems and Technology (TIST)*, *12*(6), 1-45.

Kalutharage, C. S., Liu, X., Chrysoulas, C., & Bamgboye, O. (2024, June). Neurosymbolic learning in the XAI framework for enhanced cyberattack detection with expert knowledge

integration. In *IFIP International Conference on ICT Systems Security and Privacy Protection* (pp. 236-249). Cham: Springer Nature Switzerland.

Mekrache, A., Ksentini, A., & Verikoukis, C. (2024). Machine Reasoning in FCAPS: Towards Enhanced Beyond 5G Network Management. *IEEE Communications Surveys & Tutorials*.

Forti, S., Breitenbücher, U., & Soldani, J. (2022). Trending topics in software engineering. *ACM SIGSOFT Software Engineering Notes*, *47*(3), 20-21.

Zhu, R. (2024). Automatic construction and updating of knowledge base from log data.

Beebe, N. H. (2025). A Complete Bibliography of ACM Transactions on Intelligent Systems and Technology (TIST).

Koussouris, S., Dalamagas, T., Figueiras, P., Pallis, G., Bountouni, N., Gkolemis, V., ... & Agostinho, C. (2024, June). Bridging Data and AIOps for Future AI Advancements with Human-in-the-Loop. The AI-DAPT Concept. In *2024 IEEE International Conference on Engineering, Technology, and Innovation (ICE/ITMC)* (pp. 1-8). IEEE.

Locher, M. G. (2023). Optimizing IT operations with AIOps: an investigation into the opportunities and challenges for enterprise adoption.

Remil, Y. (2023). *A data mining perspective on explainable AIOps with applications to software maintenance* (Doctoral dissertation, INSA de Lyon).

Sabharwal, N., & Bhardwaj, G. (2022). Hands-on AIOps. *Apress eBooks. https://doi.org/10.1007/978-1-4842-8267-0*.

Remil, Y., Bendimerad, A., Mathonat, R., & Kaytoue, M. (2024). Aiops solutions for incident management: Technical guidelines and a comprehensive literature review. *arXiv preprint arXiv:2404.01363*.

Paleyes, A. (2024). *Towards Maintainable and Explainable AI Systems with Dataflow* (Doctoral dissertation).

Tanikonda, A., Katragadda, S. R., Peddinti, S. R., & Pandey, B. K. (2021). Integrating AI-Driven Insights into DevOps Practices. *Journal of Science & Technology*, *2*(1).

Irfan, K., & Daniel, M. (2024). AI-Augmented DevOps: A New Paradigm in Enterprise Architecture and Cloud Management.

Oladoja, T. (2024). Artificial Intelligence-Driven Innovations in VLSI, DevOps Security, and Cloud-Native Platforms: Addressing Challenges in Modern Technology Development.

Nagmoti, N. S., Srivastava, I., & Damle, M. (2025). AI-Driven Enhancements in Cloud-Native DevOps Boosting Automation, Deployment, and Monitoring. In *Artificial Intelligence for Cloud-Native Software Engineering* (pp. 203-236). IGI Global Scientific Publishing.

Remil, Y. (2023). *A data mining perspective on explainable AIOps with applications to software maintenance* (Doctoral dissertation, INSA de Lyon).

Lyu, Y., Rajbahadur, G. K., Lin, D., Chen, B., & Jiang, Z. M. (2021). Towards a consistent interpretation of aiops models. *ACM Transactions on Software Engineering and Methodology (TOSEM)*, *31*(1), 1-38

Mulongo, N. Y. (2024, October). Key Performance Indicators of Artificial Intelligence For IT Operations (AIOPS). In *2024 International Symposium on Networks, Computers and Communications (ISNCC)* (pp. 1-8). IEEE.

Yang, L., & Rossi, D. (2022). Quality monitoring and assessment of deployed deep learning models for network AIOps. *IEEE Network*, *35*(6), 84-90.

Cheng, Q., Sahoo, D., Saha, A., Yang, W., Liu, C., Woo, G., ... & Hoi, S. C. (2023). Ai for it operations (aiops) on cloud platforms: Reviews, opportunities and challenges. *arXiv preprint arXiv:2304.04661*.

Cheng, Q., Sahoo, D., Saha, A., Yang, W., Liu, C., Woo, G., ... & Hoi, S. C. (2023). Ai for it operations (aiops) on cloud platforms: Reviews, opportunities and challenges. *arXiv preprint arXiv:2304.04661*.

Sabharwal, N., & Bhardwaj, G. (2022). Hands-on AIOps. *Apress eBooks. https://doi.org/10.1007/978-1-4842-8267-0*.

Locher, M. G. (2023). Optimizing IT operations with AIOps: an investigation into the opportunities and challenges for enterprise adoption.

Veluru, S. P. (2021). Leveraging AI and ML for Automated Incident Resolution in Cloud Infrastructure. *International Journal of Artificial Intelligence, Data Science, and Machine Learning*, *2*(2), 51-61.

Tu, S., Waqas, M., Badshah, A., Yin, M., & Abbas, G. (2023). Network intrusion detection system (NIDS) based on pseudo-siamese stacked autoencoders in fog computing. IEEE Transactions on Services Computing, 16(6), 4317-4327.

Mondru, A. K., Shreyas, R. B., & Anabathula, T. S. (2024). A Roadmap to Success: Strategies and Challenges in Adopting Aiops for it Operations. *International Journal of Interpreting Enigma Engineers (IJIEE)*, *1*(2).

Becker, S., Schmidt, F., Gulenko, A., Acker, A., & Kao, O. (2020, December). Towards aiops in edge computing environments. In *2020 IEEE International Conference on Big Data (Big Data)* (pp. 3470-3475). IEEE.

Goethals, T., Volckaert, B., & De Turck, F. (2021). Enabling and leveraging AI in the intelligent edge: A review of current trends and future directions. *IEEE Open Journal of the Communications Society*, *2*, 2311-2341.

Wang, X., Badshah, A., Tu, S., & Waqas, M. (2021, November). Blockchain-Based Security Management Platform. In 2021 2nd Asia Symposium on Signal Processing (ASSP) (pp. 118-121). IEEE.

Jagatheesaperumal, S. K., Pham, Q. V., Ruby, R., Yang, Z., Xu, C., & Zhang, Z. (2022). Explainable AI over the Internet of Things (IoT): Overview, state-of-the-art and future directions. *IEEE Open Journal of the Communications Society*, *3*, 2106-2136.

Remil, Y. (2023). *A data mining perspective on explainable AIOps with applications to software maintenance* (Doctoral dissertation, INSA de Lyon).

Polu, O. R. (2024). AI-Driven Prognostic Failure Analysis for Autonomous Resilience in Cloud Data Centers. *Journal ID*, *2563*, 4512.