*International Journal of Computational and Experimental Science and ENgineering (IJCESEN)*

*Vol. 11-No.4 (2025) pp. 9688-9696*
*http://www.ijcesen.com*

***ISSN: 2149-9144***



**Research Article**

# AI-First Reliability Engineering: Redefining SRE with Autonomous AI Agents

**Saravanan Raj***

Independent Researcher, USA
* **Corresponding Author Email:** reach.saravanan.raj@gmail.com - **ORCID:** 0000-0002-1147-9050

**Article Info:**

**DOI:** 10.22399/ijcesen.451
**Received :** 21 October 2025
**Revised :** 10 December 2025
**Accepted :** 15 December 2025

**Keywords**

AI-First Reliability Engineering,
Autonomous Agents,
AIOPS,
Incident Management Automation,
Site Reliability Engineering

**Abstract:**

Abstract should be about 100-250 words. It should be written times new roman and 10 punto. New issues in modern cloud operation are more than ever before because the complexity of the system and the scale of the system surpasses the traditional practice of manual reliability. Algorithms Site Reliability Engineering teams deal with data volumes of telemetry that are gigaworthy, with alert fatigue, and the repetitive cycle of responding to incidents that hazard and wastes engineering resources without always stopping expensive service disruptions. The appearance of artificial intelligence in the IT activity, specifically, the use of autonomous agent systems driven by large language models, allows reconsidering the concept of reliability management fundamentally. Multi-agent architectures use specialized monitoring, diagnosis and remediation components to work together to manage incidents with little human intervention. Practical applications show significant drops in mean time to resolution, drastic reduction of on-call load and quantifiable increases in system availability indicators. Organizations that apply AI-first reliability models realize strong economic payoffs via safeguarded earnings and recaptured engineering time and at the same period, lowered engineer burnout and enhanced workforce contentment. They include critical success factors such as setting up strong observability bases, deploying explainability through transparency, engages in rigorous testing of chaos engineering, and has proper human control over high stakes decisions. The transition to autonomous reliability management is a paradigm shift in which human knowledge is interested in strategic system design and AI agents are in charge of operational speed and scale, executing adopters as leaders in providing robust digital services.

## 1. Introduction

Online services today are expanding at scale and complexity that challenges the conventional approaches of reliability to its limits. Clouds are made up of thousands of micro services and millions of servers that produce a flood of telemetry that is difficult to digest without computers aiding the task. The development of artificial intelligence in IT operations is a shift in paradigm in the manner in which organizations consider the reliability of a system and the management of incidence. Based on extensive studies of the outcomes of workshops about AIOps, it is now clear that machine learning and artificial intelligence are required in the management of the complexity and scale of the modern distributed systems wherein the conventional rule-based operational concepts are no longer relevant in dealing with the volume and speed of operational data [1]. The cost and frequency of service outages have grown and organizations are experiencing immense financial and reputational losses due to system failures disseminating through interconnected microservices architectures.

In the current digital economy, the economic interests of a reliable system are beyond exceptionally high. A study on the microservices architectures has found out that distributed nature of the modern cloud systems has presented new sets of failures connected to service communication, network partitions, and cascading dependencies previously unknown to traditional monolithic systems and made it highly difficult to detect and resolve incidents [2]. The cost analysis of the industry reveals that application downtime costs an organization hundreds of thousands of dollars per hour, and in some cases, enterprises incurred multi-million-dollar losses in case of major incidents. Simultaneously, the operation of the Site Reliability

Engineering staff is more difficult than ever, drowning in an deluge of thousands of alerts per week, and devoting large percentages of their time to managing incidents instead of actively improving their systems. This alert fatigue is one of the factors that lead to burnout and decreased efficiency of engineers in distinguishing the true serious cases among the clutter of false positives and low priority alerts.

The core issue underlying the contemporary operations teams is an increasing gap in the level of reliability whereby complex failures are happening at a rate exceeding the ability of humans to perceive, diagnose, and amend them. Slow incident response is more than loss of immediate revenue to include broken customer trust, tarnished brand reputation, and market side effects as competitors leverage their dependability in services as a distinguishing factor. This paper suggests an AI-First SRE model that deploys intelligent autonomous agents across the incident lifecycle, and uses new developments around large language models and multi-agent systems to develop collaborative machine responders in the context of the speed and scale requirements of modern cloud infrastructure.

## 2. Case Study Background and Use Case Overview.

Application of AI-first reliability engineering principles may be exemplified by analyzing a big cloud services company that initiated a massive automation-first reliability program. This organization had a worldwide cloud platform that served millions of users in various geographical locations with hundreds of microservices that produced tremendous amounts of telemetry information each day. The difficulty of scaling such infrastructure is indicative of more general industry trends recorded in serverless computing literature, which established that large cloud vendors handle billions of invocations of functions per day on millions of unique applications, and that the workload characteristics are largely varied in terms of executing patterns, resource demands, and failure modes that are unsuitable to handle with simple rule-based management strategies [3]. The conventional Site Reliability Engineering team model was not sufficient to manage the pace and complexity of incidents that were taking place in this distributed system environment.

Preliminary evaluation of the work environment shows that there are systematic inefficiency in the incident management procedures. A significant percentage of incidences were recurrent styles that had been experienced severally before, but every time the incidence took place, it had to be computer-investigated and fixed by the available engineers. Studies of software failures in cloud computing systems have reported that the patterns of failures tend to be highly time- and spatially correlated, with some categories of bugs occurring repeatedly in many services and components due to the existence of common infrastructure dependencies and distributed system design antipatterns [4]. The average amount of time that engineers spent on incident response operations occupied major parts of the weekly work schedules, and especially the on-call rotations that mandated immediate response at any hour of the week or night. Complex outages often necessitated the gathering of expertise across many different teams leading to coordination overhead and a delay in finding a resolution.

The scale problems that this organization dealt with were representative of larger issues in the industry that involved manual incident management strategies. Infrastructure faults, on minor scale, were frequently causing cascading alert storms with dozens of correlated alerts, which is far too many to easily spot root causes within the noise of human operators. Multi-failure mode complex outages involved a lot of investigation of logs, metrics, and traces that were distributed across a variety of monitoring systems and dashboards. The company management realized that the further linear expansion of the human SRE team by the infrastructure could not be economically viable and would not be effective. The concept of the change to an AI-first reliability model involved the deployment of autonomous machine agents that will be able to work 24/7 to manage common patterns of incidents but only escalate when it comes to new or high-stake situations that require human intervention.

The strategic goals of this transformation program involved quantitative performance change as well as qualitative change in terms of work experience in the engineering profession. In particular, the company was aiming at tremendous improvements in the mean time to resolution of generic types of incidents with the help of automated diagnosis and remedication features. The other objective, which was as vital, was to reclaim engineering hours of repetitive fire fighting and have the SRE team concentrate on proactive reliability efforts that would include architecture improvements, capacity planning and automation development. The metrics of success would extend beyond the conventional indicators of reliability such as uptime percentage and the number of incidents but also cover human metrics such as the reduction of on-call burden, the satisfaction levels of the engineers, and retention

rates of the team which would indicate the sustainability of operational practices.

## 3. Implementation Architecture and Methodology.

To ensure the AI-first SRE system employs the right balance between autonomy and safety, as well as to achieve the effective distribution of tasks between specialized agent components, the technical implementation of the system needed careful design. The adopted strategy followed a multi-agent platform in which various agents had specialised knowledge in specific fields of operation, and they interacted with a coordinating supervisor layer. This design philosophy follows reported trends in artificial intelligence about IT operations, in which heterogeneous data such as logs, metrics, traces, and topology data need to be integrated to develop overall insights into the state of the system and failure modes [5]. The implementation process was carried out in several phases that were accomplished in a long timeline with each stage being based on the lessons learned in the earlier stage and also taking into account the ongoing feedback in the process of operational deployment.

The initial stage was aimed at the development of a strong observability base able to record extensive telemetry of all the system elements with a high degree of granularity and timeliness to make real-time decisions under the guidance of AI agents. The organization invested in the standardization of open telemetry frameworks to guarantee a consistent instrumentation in a wide range of microservices deployed in any type of programming language and any framework. This single platform of observability would handle huge amounts of operational data with low latency between event occurrence and availability to analyse and attain high data completeness rates, which is essential to accurate AI inference. The first observability agent was introduced, which incessantly scanned telemetry streams and identified anomalies or threshold violations instead of numerous fixed threshold-based notifications with a dynamic anomaly detector that adjusted to changing baseline trends and minimized false positives by a significant margin.

The fundamental multi-agent system design consisted of four domain agents that would be domain specific agents, each driven by large language models that have been fine-tuned on large volumes of historical event data and operational manuals. An agent of Kubernetes infrastructure that is aimed to identify the problems that are associated with container orchestration, scheduling of pods, resource limits, and node health issues. An application logs agent that is capable of processing log streams with the aim of identifying error patterns, deriving useful diagnostic data, and correlating events between distributed services. A performance metrics agent was used to monitor thousands of different key performance indicators to identify irregularities in the latency and throughput, resource utilization, and additional quantifiable indicators of system health. A runbook agent was operational at any given moment to allow access to documented remediation procedures, troubleshooting guides and architectural knowledge to access contextually specific information on responding to the incident.

Those specialized agents acted under the control of a supervisor agent who managed to coordinate the investigation processes in the response to incidents. In the case of incidents, the supervisor created the investigation plans and distributed parallel tasks to the corresponding specialist agents according to preliminary symptoms and the context of the system. The analysis of root causes in multitier cloud services has shown evidence that adequate diagnosis necessitates causal reasoning in the bonds between noticed symptoms and the actual root causes which are situated beneath faults [6]. The supervisor agent took the findings of specialist agents and built coherent stories of what went wrong and why with sophisticated reasoning and causal inference to prevent confusing correlation with causation in complicated distributed failure cases.

The autonomy remedial functions were set in place cautiously in terms of safety guard rails and risk management. The organization coded into a huge library of verified remediation measures by using a large amount of chaos engineering experiments that systematically tested failure injection and recovery processes. All the actions were rated by the degree of risk on the basis of the possible characteristics of the blast radius and reversibility. Service restarts, configuration rollbacks or resource scaling, which are low-risk operations, could be autonomously performed by AI agents when their confidence scores rose above set limits. Riskier and new situations beyond the action space training necessitated the workload to be escalated to human engineers to be approved and supervised. This human-in-the-loop architecture was such that automation could increase system reliability instead of compromising it, eliminating the possibility of cases where autonomous behavior might be unintentionally harmful to an incident or create new failure modes.

The mechanism of the system included the continuous learning process that enhanced

performance throughout the time due to the performance through the operational experience. All automated and manual incident resolutions were stored in an organized knowledge base by vector embedding techniques that facilitated searches and pattern matching through semantics across millions of past case resolutions. The AI agents showed the quantifiable improvement of accuracy in the following months of use as the training data was increased, and the model improvements reflected the evident limitations. The collected feedback by the on-call engineers brought important indications regarding the cases when AI recommendations were not right or not complete, which prompted the iterative process of improving the prompts, reasoning logic, and action policies. Specific emphasis was placed on reducing the risks of hallucinations of large language models with methods such as multi-agent consensus validation, evidence-linking conditions, and restrictions to inferences based on evidence-supported conclusions as opposed to hypothetical explanations.

## 4. Results and Performance Outcomes.

The AI-first reliability system production implementation resulted in quantifiable advances on various levels of operational performance assessed throughout a long period of evaluation. Mean time to resolve experienced dramatic decreases over the baseline manual procedures with incidents that previously took hours of research and remediation being solved within minutes by automated diagnosis and response. Studies of AIOps to manage incident have reported that companies that adopted state-of-the-art AI methods recorded significant MTTR improvements (ranging between seventy and eighty-five percent) in the most frequent type of incidents, and an especially positive outcome with repeating groups of failures that can be easily identified by AI methods and remedied [7]. This increase in speed directly resulted in the improved availability of the service as at a faster pace, the incident resolution time shortened, as well as, it allowed to achieve more demanding service level targets.

MTTR distribution was also different among the different categories of incidence depending on the complexity and novelty of failure modes. Well-known types of failures like database connection exhaustion, memory leaks, or deployment-related regressions also demonstrated the most radical improvements as automated systems ran known remediation steps in seconds as opposed to the minutes or hours to investigate and take action of failures when a human was involved. More complex events that are multiple interacting failures or new cases not covered by the training distribution still demanded significant human intervention, although even here the AI agents were useful as they quickly synthesized the available information and formulated the initial hypotheses which could be investigated by humans [8]. The general system availability measures were raised significantly, which translated to a significant decrease in the number of hours of downtimes per year and a relation to revenue and consumer satisfaction protection respectively.

In addition to the speed of resolution, the AI system radically lowered the operational workload of human engineers by automating repetitive incident response processes. The agents were able to cope with a very high percentage of incidences without any human input, both on the first detection and identification of the root cause and remediation. Tens of thousands of engineering hours each year were reclaimed through this automation coverage which would have been otherwise used in manual firefighting efforts. Acute care engineers indicated that their overall incident response time commitment as an on-call decreased significantly in weekly incidents, and in particular, the overnight and weekend pages that interrupted personal time and led to burnout decreased significantly. This productivity gain was of an economic nature as it involved direct savings in labor costs as well as indirect benefits in terms of time saved in allowing engineers to focus on active reliability improvements, building improvements, and the expansion of automation that yielded compounding returns over time.

The reliability increases reflected in greater adherence to service level goals in many measures that are monitored by the organization. The objectives of API response time were achieved more regularly because the incidents that led to the degradation of latency were identified and addressed faster. Jobs of processing of data were executed faster under the service level agreements because of the speed with which the pipeline failures and resource allocation were resolved. Above all, the number of customer-facing incidents fell significantly as the AI system was found to be efficient at identifying and resolving them before they became an issue of concern to the users. There were corresponding customer satisfaction metrics, as the reliability perception scores were improved and the number of customer-reported incident complaints was reduced. These reliability gains were measured in business terms as the amount of revenue saved through downtime avoidance, the customer retention rates that could be attributed to a better reputation for reliability, and competitive

advantages in markets where availability of services was used as a major differentiating factor.

The human factors outcomes were also important on the sustainability of the organization. On-call engineer burnout indicators improved significantly as determined by the validated psychological assessment tools and the scores dropped to worrying scores, which indicate high stress levels and exhaustion to healthier scores. Employee satisfaction surveys showed a high level of positive emotion in relation to improvement of work-life balance, ability to get down to significant engineering tasks as opposed to menial labor as well as the confidence that the organization can be reliable without having to work wonders on a single head. The level of voluntary attrition in the SRE organization decreased significantly, saving both humanity and resources on expensive attrition and maintaining institutional knowledge and expertise. The distribution of time changed significantly toward incidence response in reaction to incidents and strategic work such as capacity planning, system architecture enhancements and creation of other automation facilities. These cultural and organizational advantages affirmed the business case of AI-first reliability beyond pure technical performance metrics, that investments in operational AI capabilities brought returns in terms of better talent retention and engineering effectiveness.

## 5. Lessons Learned and Future Implications.

The experience of the process of introducing AI-first reliability engineering brought many lessons that can be applied in any organization seeking to introduce a similar change to their operational processes. One of the most important lessons was the need to implement the phased rollout strategies which would instill trust, as the simplified autonomous authority will be proven correct. The first implementation was in recommend-only shadow mode, which enabled the verification of the quality of AI agent decisions against the human expert one without taking risks of wrong automated operations. Studies into intelligent operations incorporating large language models have emphasized how proper trust boundaries can only be set through delicate calibration of automation benefits and possible failure mode with cautious expansion of autonomous capabilities as models demonstrate themselves trustworthy in a production environment [9]. Companies must avoid the urge to go all the way to automation at once, but instead take staged steps, which gain the trust of the stakeholders due to a steady track record of success over a prolonged duration.

Explainability and transparency appeared to be key criteria towards organizational acceptance and successful human-AI cooperation. Engineers were significantly more willing to believe and take action on AI recommendations when the system could clearly articulate the reasoning of what evidence informed the conclusions and what logic linked observations to the diagnoses. Every incident analysis contained formal descriptions of how the recommendations were related to particular supporting information like log or coded errors, measurements that had anomalies, or previous incidences illustrating the identical patterns. This explainability had two functions of instilling confidence among operators in the quality of decisions made by AI and learning, under which engineers were able to know not just what to do but also why that action was performed on root causes. To ensure that black box systems are adopted, organizations that are applying AI operations system should focus on explainability features, even at the expense of model complexity or inference speed, despite their accuracy.

Reducing the risk of hallucinations and false positive situations demanded continuous consideration in the form of systematic testing and model optimization. Even early deployment exposed that large language models could sometimes give plausible-sounding but false-fact explanations or advice, especially in failure cases not in the training distribution. The use of multi-agent consensus mechanisms where other independent agents reviewed incidents and compared their findings was used to raise cases where individual agents made different conclusions that led to further investigation. Studies of causal graph-based root cause analysis methods have shown that root cause diagnosis accuracy can be significantly enhanced by the addition of designed knowledge regarding system dependencies and causal dependencies and can greatly exceed that of pure pattern matching on symptoms [10]. In order to make AI inference based on proven system models, organizations are advised to integrate statistical machine learning with domain knowledge encoding as well as causal reasoning abilities instead of being reliant on the learned correlations only.

The human expertise-AI capabilities partnership model was more effective than considering automation as the substitute of the human operator. Studies of complex incidences showed that human judgment was still necessary in situations that involved uncertain evidence, security concerns, data integrity, or business situation concerns other than pure technical system condition. The best allocation of roles put AI agents in a position to

manage speed and scale dimensions such as synthesizing data quickly, matching patterns on historical cases and performing routine remediation process but retained human supervision in case of contextual judgment, innovative problem-solving to new failures, and responsibility in making decisions with a notable business impact. The companies must position the organization of AI operations programs as an augmentation that empowers human engineers to a supervisory position and a strategic position than a displacement, to capture the technical realities and respond to the anxieties of automation among employees.

On the longer-term implications, when AI-first reliability practices were successfully implemented, it is possible to expect fundamental change in the way organizations deal with operational engineering. The market trend shows a vast increase in the use of AIOps platform and the world market is expected to increase exponentially within the next few years due to the growing complexity of cloud-native systems and the proven worth of the solution by early adopters [4]. With the ongoing development of AI agents capabilities due to the growing sophistication of language models, reasoning engines, and methods of causal inference, companies can expect the evolution of more autonomous self-healing systems with not only reactive incident response, but also proactive prediction and prevention of failures before user impact. Best practices and standards will inevitably change to meet testing needs on AI system modifications, audit processes of automated decisions, and data sharing systems to enhance models with aggregate learnings across companies. The future is clearly defined by the trajectory of human-AI cooperation making reliability engineering a standard practice, and the role of the SRE professionals will shift to that of firefighters to developers of even more intelligent autonomous operational systems.

***Table 1:*** *Baseline Infrastructure and Incident Characteristics [3, 4]*

| Metric Category | Description | Baseline State |
|---|---|---|
| Infrastructure Scale | Geographic regions, microservices, and server instances | Global platform with extensive distributed architecture |
| Daily Telemetry Volume | Log entries, metric points, and trace spans generated | Terabytes of operational data requiring processing |
| Monthly Incident Volume | Total incidents handled by engineering teams | Thousands of incidents across all severity levels |
| Engineer Team Size | Number of SRE professionals managing operations | Dedicated reliability engineering organization |
| Incident Response Time | Average time engineers spent on incident activities weekly | Substantial portion of engineering capacity consumed |
| Repetitive Incident Rate | Proportion of recurring failure patterns | Majority showed patterns from previous occurrences |
| War Room Requirements | Teams and duration for complex outages | Multiple teams coordinating for extended periods |

***Table 2:*** *AI Agent Specializations and Responsibilities [5, 6]*

| Agent Type | Primary Function | Data Sources | Key Capabilities |
|---|---|---|---|
| Observability Agent | Continuous monitoring and anomaly detection | Unified telemetry streams across all services | Real-time detection with reduced false positives |
| Kubernetes Infrastructure Agent | Container orchestration and node health | Pod scheduling, resource constraints, cluster state | Infrastructure-specific failure pattern recognition |
| Application Logs Agent | Error pattern identification and event correlation | Service logs across multiple frameworks | Cross-service event correlation and diagnostics |
| Performance Metrics Agent | Key performance indicator monitoring | Latency, throughput, resource utilization metrics | Dynamic anomaly detection across thousands of KPIs |
| Operational Runbook Agent | Knowledge retrieval and procedure access | Documentation, troubleshooting guides, architecture | Contextual information retrieval for remediation |
| Supervisor Agent | Investigation orchestration and synthesis | Aggregated findings from specialist agents | Causal reasoning and narrative construction |
| Remediation Agent | Corrective action execution | Validated action library and confidence scores | Risk-aware autonomous execution with escalation |

***Table 3:*** *Operational Metrics Before and After AI Implementation [7, 8]*

| Performance Dimension | Metric Description | Direction of Improvement |
|---|---|---|

| | | |
|---|---|---|
| Mean Time to Resolution | Median incident resolution duration | Substantial reduction from hours to minutes |
| Incident Automation Coverage | Percentage of incidents handled autonomously | Majority of incidents resolved without human intervention |
| Engineering Hours Reclaimed | Annual reduction in manual incident response time | Tens of thousands of hours redirected to strategic work |
| Weekly On-Call Burden | Average incident response time per engineer | Dramatic decrease in weekly commitment hours |
| Night-Time Pages | Overnight alerts requiring engineer wake-up | Significant reduction in after-hours disruptions |
| System Availability | Overall uptime percentage | Measurable improvement in service reliability |
| Customer-Facing Incidents | Quarterly count of user-impacting events | Substantial decrease in customer impact frequency |
| SLO Compliance | Achievement of service level objectives | Improved consistency across multiple metrics |
| Engineer Burnout Scores | Psychological assessment of stress levels | Notable improvement toward healthier ranges |
| Team Attrition Rate | Annual voluntary departure percentage | Reduction to more sustainable retention levels |

***Table 4:*** *Critical Elements for AI-First Reliability Adoption [9, 10]*

| Success Factor Category | Key Elements | Implementation Considerations |
|---|---|---|
| Phased Rollout Strategy | Shadow mode validation before autonomous operation | Incremental authority expansion based on proven accuracy |
| Observability Foundation | Comprehensive telemetry with minimal latency | Standardized instrumentation across diverse services |
| Transparency Mechanisms | Structured explanations with evidence linking | Clear reasoning for building operator trust |
| Hallucination Mitigation | Multi-agent consensus and evidence constraints | Systematic testing through chaos engineering |
| Risk Management | Action classification and confidence thresholds | Human oversight for high-stakes decisions |
| Continuous Learning | Knowledge base updates and feedback loops | Model refinement based on operational experience |
| Cross-Functional Teams | Combined SRE and machine learning expertise | Balanced technical and operational perspectives |
| Change Management | Workforce engagement and training programs | Positioning AI as augmentation rather than replacement |
| Testing Rigor | Validation across diverse failure scenarios | Chaos engineering to verify autonomous behavior |
| Human-AI Partnership | Clear responsibility boundaries | Strategic human judgment with AI speed and scale |

## 6. Conclusions

AI-first reliability engineering is a disruptive model of enabling significantly better uptime, operational performance and engineering sustainability in contemporary cloud infrastructure. The pieces of evidence provided by studying the case in detail prove that well-planned autonomous AI agent systems can save a significant difference in the mean time to resolve, reclaim thousands of hours of engineering time spent on incident response processes, and increase system availability indicators with a high return on investment in the relatively limited time frames. These quantitative performance pluses are complemented by qualitative gains that are equally substantial, such as a decrease in burnout levels in engineers, an increase in work/ life balance,an increase in job satisfaction due to emphasis on creative engineering work and lower rate of attrition that

increases organizational capacity and knowledge retention. To stay competitive in the provision of reliable digital services, organizations ought to consider the extent to which autonomous AI agents can tackle particular operational pain points and reliability issues, starting with full-scale reliability tests to understand the common patterns of incidents that can initially be automated as part of their initial automation pilot. Critical success factors are not limited to technical implementation but include management of organizational change, which entails visible transparency through AI reasoning that can be understood and read by human operators, rigorous testing using chaos engineering to ensure autonomous behavior effectiveness in multi-failure situations, and gradual rollout, which creates trust by demonstrating the accuracy of autonomous behavior, before more autonomous authority is granted. The benefits of an AI-first reliability practice implementation are large-scale and multidimensional, and less downtime is an opportunity to protect revenue, reallocated engineering time can be used to concentrate on the positive reaction on the system, and the well-being of engineers leads to a high team morale and enhanced organizational performance. The future of extremely accessible services will rely upon the principles of cooperative alliances between human specialist knowledge and independent AI-based functions and utilize synergistic advantages of machine pace and volume in association with human view and originality. Those that start practicing AI-first reliability today, based on the successful examples of others and building on the previous experience, will be in the position to be the first to achieve the level of reliability standards that seemed impossible a few years ago and will continue to enhance the experience of the engineering teams that need to maintain the critical infrastructure.AI is applied in different fields and reported [11-26].

## Author Statements:

- **Ethical approval:** The conducted research is not related to either human or animal use.
- **Conflict of interest:** The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper
- **Acknowledgement:** The authors declare that they have nobody or no-company to acknowledge.
- **Author contributions:** The authors declare that they have equal right on this paper.
- **Funding information:** The authors declare that there is no funding to be acknowledged.
- **Data availability statement:** The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

## References

[1] Jasmin Bogatinovski et al., "Artificial Intelligence for IT Operations (AIOPS) Workshop White Paper," arXiv preprint arXiv:2101.06054, 2021. [Online]. Available: https://arxiv.org/abs/2101.06054

[2] J. Soldani, D. A. Tamburri, and W. J. Van Den Heuvel, "The pains and gains of microservices: A Systematic grey literature review," Journal of Systems and Software, 2018. [Online]. Available: https://www.sciencedirect.com/science/article/abs/pii/S0164121218302139

[3] Youcef Remil, "AIOps Solutions for Incident Management: Technical Guidelines and A Comprehensive Literature Review," arXiv preprint arXiv:2404.01363, 2024. [Online]. Available: https://arxiv.org/abs/2404.01363

[4] MarketsandMarkets, "AlOps Platform Market by Offering (Platforms (Domain-centric, Domain-agnostic), Services (Professional, Managed)), Application (Infrastructure Management, ITSM, Security & Event Management), Deployment Mode, Vertical and Region - Global Forecast to 2028," 2023. [Online]. Available: https://www.marketsandmarkets.com/Market-Reports/aiops-platform-market-251128836.html

[5] Britta, "ITSM outcomes and how to measure them," 2024. [Online]. Available: https://www.servicenow.com/community/itsm-articles/itsm-outcomes-amp-how-to-measure-them/ta-p/2309407

[6] Mohammad Shahrad et al., "Serverless in the Wild: Characterizing and Optimizing the Serverless Workload at a Large Cloud Provider," in *Proc. USENIX Annual Technical Conference (ATC)*, 2020, pp. 205-218. [Online]. Available: https://www.usenix.org/conference/atc20/presentation/shahrad

[7] Domenico Cotroneo et al., How bad can a bug get? an empirical analysis of software failures in the OpenStack cloud computing platform," ESEC/FSE 2019: Proceedings of the 2019 27th ACM Joint Meeting on European Software Engineering Conference and Symposium on the Foundations of Software Engineering, 2019. [Online]. Available: https://dl.acm.org/doi/10.1145/3338906.3338916

[8] CrowdStrike, "Automated Root Cause Analysis". [Online]. Available: https://www.datadoghq.com/dg/apm/root-cause-analysis/

[9] Lingzhe Zhang et al., "A Survey of AIOps in the Era of Large Language Models," ACM Computing Surveys, 2025. [Online]. Available: https://dl.acm.org/doi/10.1145/3746635

[10] Meng Ma et al., "AutoMAP: Diagnose Your Microservice-based Web Applications Automatically," WWW '20: Proceedings of The Web Conference 2020, 2020. [Online]. Available: https://dl.acm.org/doi/10.1145/3366423.3380111

[11]Madhavi Mangalarapua. (2025). Evaluation of DNA damage and repair in Radiographers and Dental Surgeons using X-ray machines in Dental Clinics. *International Journal of Natural-Applied Sciences and Engineering*, *3*(1). https://doi.org/10.22399/ijnasen.14

[12]Ibeh, C. V., & Adegbola, A. (2025). AI and Machine Learning for Sustainable Energy: Predictive Modelling, Optimization and Socioeconomic Impact In The USA. *International Journal of Applied Sciences and Radiation Research* , *2*(1). https://doi.org/10.22399/ijasrar.19

[13]Soyal, H., & Canpolat, M. (2025). Intersections of Ergonomics and Radiation Safety in Interventional Radiology. *International Journal of Sustainable Science and Technology*, *3*(1). https://doi.org/10.22399/ijsusat.12

[14]Ankit, & Amritpal Singh. (2025). Optimized Architecture for Efficient VM Allocation and Migration in Cloud Environments. *International Journal of Computational and Experimental Science and Engineering*, *11*(2). https://doi.org/10.22399/ijcesen.1466

[15]García, R. (2025). Optimization in the Geometric Design of Solar Collectors Using Generative AI Models (GANs). *International Journal of Applied Sciences and Radiation Research* , *2*(1). https://doi.org/10.22399/ijasrar.32

[16]Vishwanath Pradeep Bodduluri. (2025). Social Media Addiction and Its Overlay with Mental Disorders: A Neurobiological Approach to the Brain Subregions Involved. *International Journal of Sustainable Science and Technology*, *3*(1). https://doi.org/10.22399/ijsusat.3

[17]Ujjwal Raj. (2025). The Serverless Paradigm: Abstraction, Elasticity, and Event-Driven Computing in Modern Cloud Architectures. *International Journal of Computational and Experimental Science and Engineering*, *11*(4). https://doi.org/10.22399/ijcesen.4088

[18]Harsha Patil, Vikas Mahandule, Rutuja Katale, & Shamal Ambalkar. (2025). Leveraging Machine Learning Analytics for Intelligent Transport System Optimization in Smart Cities. *International Journal of Applied Sciences and Radiation Research* , *2*(1). https://doi.org/10.22399/ijasrar.38

[19]Jhansi Rani Ganapa, Poonam Joshi, T Amitha, Sandip Rahane, N. Ravinder, Jignesh Jani, … Chandreshkumar Vyas. (2025). Security and Privacy Challenges in Deep Learning Models Hosted on Cloud Platforms. *International Journal of Computational and Experimental Science and Engineering*, *11*(3). https://doi.org/10.22399/ijcesen.3235

[20]Chui, K. T. (2025). Artificial Intelligence in Energy Sustainability: Predicting, Analyzing, and Optimizing Consumption Trends. *International Journal of Sustainable Science and Technology*, *3*(1). https://doi.org/10.22399/ijsusat.1

[21]V. Ananthakrishna, & Chandra Shekhar Yadav. (2025). QP-ChainSZKP: A Quantum-Proof Blockchain Framework for Scalable and Secure Cloud Applications. *International Journal of Computational and Experimental Science and Engineering*, *11*(1). https://doi.org/10.22399/ijcesen.718

[22]Madane, S., Kamble, V., & Chavan, G. (2025). Cyber Chain – Merging Blockchain with Cyber Security. *International Journal of Applied Sciences and Radiation Research* , *2*(1). https://doi.org/10.22399/ijasrar.42

[23]Kumari, S. (2025). Machine Learning Applications in Cryptocurrency: Detection, Prediction, and Behavioral Analysis of Bitcoin Market and Scam Activities in the USA. *International Journal of Sustainable Science and Technology*, *3*(1). https://doi.org/10.22399/ijsusat.8

[24]Olola, T. M., & Olatunde, T. I. (2025). Artificial Intelligence in Financial and Supply Chain Optimization: Predictive Analytics for Business Growth and Market Stability in The USA. *International Journal of Applied Sciences and Radiation Research* , *2*(1). https://doi.org/10.22399/ijasrar.18

[25]Fabiano de Abreu Agrela Rodrigues, & Flávio Henrique dos Santos Nascimento. (2025). Neurobiology of perfectionism. *International Journal of Sustainable Science and Technology*, *3*(1). https://doi.org/10.22399/ijsusat.6

[26]S. Jagan, B. Girirajan, Manisha Bhimrao Mane, R B, H. J., Mariam Anil, & M. Thillai Rani. (2025). Adaptive Quantum AI Models for Accelerating Deep Learning in Decentralized Cloud Architectures. *International Journal of Computational and Experimental Science and Engineering*, *11*(3). https://doi.org/10.22399/ijcesen.2493