



(REVIEW ARTICLE)



Self-Healing Software Systems: AI-Driven Fault Prediction and Recovery

Pratyosh Desaraju *

University of Central Missouri, USA.

World Journal of Advanced Engineering Technology and Sciences, 2026, 18(03), 064–071

Publication history: Received on 11 January 2026; revised on 01 March 2026; accepted on 02 March 2026

Article DOI: <https://doi.org/10.30574/wjaets.2026.18.3.0120>

Abstract

The issue with modern software systems is that they are becoming extremely complicated and interdependent, which requires fault detection and recovery as key factors for reliability and performance, especially in mission-critical settings. Self-healing computer systems are a new technology that has entered the field as a groundbreaking approach with the ability to automatically monitor, predict, and fix failures on their own. The paper reviews an approach to integrating Artificial Intelligence (AI) in self-healing architectures, specifically through machine learning, semantic reasoning, and bio-inspired strategies to enhance resilience and operational efficiency.

The paper discusses frameworks for fault prediction, anomaly detection, and automated recovery in different computing systems, such as cloud infrastructures, IoT ecosystems, and healthcare systems, based on recent studies. It also discusses how AI is now being utilized in software testing, log analysis, and cognitive fault diagnosis. The findings indicate that AI not only increases the accuracy of fault management but also delivers predictive and adaptive strategies that outperform traditional reactive models. The review concludes by describing existing challenges and future research directions required to develop reliable self-healing software systems that can be scaled to new levels.

Keywords: Self-Healing Software; AI-based Fault Prediction; Autonomous Recovery Systems; Anomaly Detection.

1. Introduction

The nature and scale of modern software systems, particularly those found in distributed, cloud-based, and IoT-enabled systems, has necessitated the development of mechanisms that can automatically detect, diagnose, and rectify errors without the involvement of humans. This aspect, referred to as self-healing, has been expedited by the introduction of Artificial Intelligence (AI), which supplements outdated approaches to surveillance and recovery with predictive analytics, machine learning algorithms, and real-time modification. It is not just that self-healing systems react to failures; they anticipate them and take corrective actions before they occur, and this ensures greater availability, resilience, and efficiency. Fault prediction and recovery systems that are based on AI are already becoming very important in mission-critical systems such as defense, healthcare, financial, and power systems, where a single failure can result in immense loss or damage.

The development of self-healing architectures is based on the convergence of different fields, such as AI, systems engineering, and software reliability. Self-healing entails the underlying mechanisms of fault detection, diagnosis, prediction, and autonomous remediation. The application of AI models at all of these stages to provide timely and accurate responses is also being implemented. The hybridization of AI and self-healing represents a step toward a paradigm shift from reactive maintenance to active and autonomous software management. The current paper systematically reviews the available literature and developments in the field of AI-based self-healing software systems by describing the frameworks, approaches, and applications that characterize this emerging field.

* Corresponding author: Pratyosh Desaraju.

2. AI Frameworks for Fault Detection and Self-Healing

Recent literature has emphasized the importance of AI-based infrastructures in supporting self-healing in both distributed and mission-critical software systems. These models usually include fault predictors, real-time anomaly detectors, and decision processors that may take automatic recovery actions. A particularly vivid example is a multi-layered, machine learning-driven architecture in which machine learning models monitor the health of distributed components, detect anomalies using predefined behavioral baselines, and implement remedial actions, such as rebooting components or switching to redundant modules [1]. This approach is centered on the use of explainable AI to increase transparency and confidence in high-safety domains by making stakeholders aware of the rationale behind fault-handling decisions.

Deep learning has been identified as having significant potential in the analysis of time-series faults in log files and telemetry data. The most critical feature of such systems is that they need to operate in dynamic environments, where the state of the system frequently evolves. Continuous learning mechanisms update the AI models in real time, allowing the system to adapt to new patterns and minimize false positives. These capabilities are critical for achieving resilience in heterogeneous and distributed environments that cannot be effectively supported by traditional fault management tools [1].

3. Predictive Maintenance in Software Systems and Cyber-Physical Environments

The self-healing capabilities have also been extended not only to conventional software environments but also to cyber-physical systems, such as power grids and industrial automation. In this case, predictive maintenance models are used to forecast component failures based on sensor data, usage logs, and environmental conditions. This data is fed through AI algorithms, which attempt to predict potential failures and either recommend or take proactive actions to mitigate them. Specifically, the concept of the digital twin, a virtual mirror of physical systems, has served as a significant driving force. It enables real-time synchronization between the physical system and its digital counterpart, allowing abundant data to be used to train AI models and test recovery strategies without interfering with real-world operations [2].

It has been demonstrated that recurrent neural networks and ensemble learning methods significantly increase fault detection rates and improve maintenance scheduling. Trained on historical datasets and continuously updated in real time, these models provide more accurate and efficient predictions. The resultant effect is a software environment capable not only of detecting current anomalies but also of foreseeing patterns that indicate imminent failures [2].

Such systems are particularly relevant in large-scale infrastructures, where a single point of failure can be disastrous. Predictive analytics are used to reduce downtime, enhance safety, and optimize resource usage. Furthermore, self-healing grid systems that incorporate AI can transform energy distribution by ensuring continuity even under unfavorable conditions [2].

4. Self-Healing in IoT Ecosystems

As more and more Internet of Things (IoT) networks are implemented in industries such as healthcare, smart homes, transportation, and agriculture, it is becoming challenging to ensure their reliability because these networks are characterized by heterogeneity, scale, and intermittent connectivity. AI is particularly relevant to self-healing and fault prediction in the context of IoT ecosystems because manual control is impractical.

Both centralized and decentralized decision-making approaches are used in self-healing IoT networks. Local intelligence enables edge devices to detect and fix faults autonomously, while centralized AI models consider global, system-wide healing plans. Machine learning-based algorithms are used to monitor communications, device health, and interactions with the environment, predict faults in advance, and carry out local recovery operations [3].

In smart agriculture systems, remote field sensors require the ability to continue functioning without supervision. The data supplied by these sensors is used by AI models to diagnose abnormal behavior, such as data stagnation or communication failures. When such issues are detected, automatic self-healing mechanisms are implemented, such as resetting devices, reloading firmware, or switching to backup modules. These autonomous recovery mechanisms are required to ensure system availability, especially in environments where connectivity is not always available or reliable [3].

This field is also beginning to implement reinforcement learning, which allows IoT systems to learn optimal healing behaviors based on the outcomes of previous attempts. Over time, these systems become more effective in handling a wide range of fault scenarios, such as temporary glitches and hardware failures [3].

5. Cloud-Based Self-Healing and Automated Recovery

The dynamism of cloud systems, their scalability, and distributed architecture are well suited to AI-driven self-healing deployment. The volume of services and microservices hosted on a cloud platform is vast, and every service or microservice comes with its own dependencies and performance metrics. By handling these services manually, it is no longer possible to guarantee their well-being or intervene by making appropriate decisions. Improved self-healing in cloud computing through AI augmentation relies on telemetry data, service logs, and performance metrics to dynamically monitor system health.

A generic architecture may be defined with three primary layers: a data collection layer (the monitoring layer), an AI-based analysis layer (to identify anomalies), and a remediation layer (to initiate automated recovery actions). These actions may include service restarts, resource redirection, service rollback, or the deployment of new service instances. AI models use confidence scores as the basis for recovery decisions, representing the probability that a particular action will be successful [4].

Moreover, cloud-based self-healing systems are built on the principles of containerization and orchestration platforms, such as Kubernetes, to manage service life cycles. They become far more responsive to service failures or degradations because AI models are integrated with the orchestration environment to trigger self-healing workflows. For example, when a containerized service begins reporting latency issues, the AI model can identify the anomaly, isolate the service, and redeploy it on a healthy node without human intervention [4].

The integration of AI with cloud orchestration improves fault tolerance and supports continuous deployment practices, as the system can cope with faults that occur during frequent updates. This enhances the overall agility and robustness of cloud-managed software development pipelines [4].

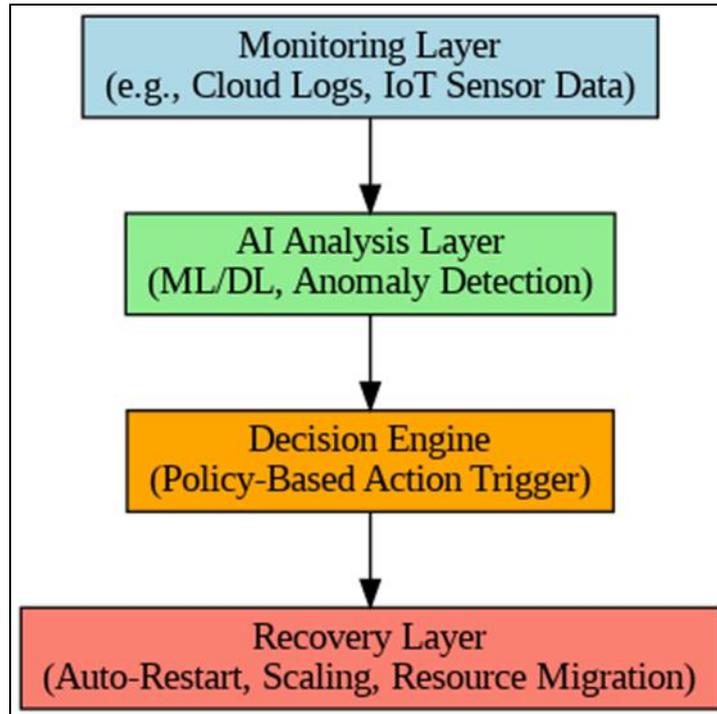
6. AI Monitoring in Critical Healthcare Applications

Malfunctions in software can cause life-threatening outcomes in mission-critical domains such as healthcare. For this reason, AI-driven self-healing functionality is particularly significant in this setting. In healthcare systems, constant availability, accuracy, and timeliness are critical requirements for software systems. To ensure reliability, AI-based monitoring systems can detect anomalies in system behavior, patterns of user interaction, and data processing pipelines.

In intensive care unit (ICU) environments, AI processes patient data in real time and operates within software platforms to generate warnings and control medical equipment. Failure of these systems can lead to incorrect diagnoses or treatments. To mitigate these risks, software performance and network communication between devices can be monitored using AI-based anomaly detection tools. When a self-healing system detects abnormal latency, memory leaks, or service unavailability, it attempts to recover by restarting services or relocating them to backup nodes [5].

Furthermore, predictive analytics helps prevent failures before they occur. Such systems can forecast resource depletion or service degradation by analyzing trends in software performance and initiating preventive measures. This type of preemptive recovery significantly reduces the risk of system downtime, which is a critical concern in healthcare contexts where every second counts [5].

These systems are maintained through continuous learning and feedback mechanisms that keep AI models informed about new operational data and emerging error patterns. In addition, rigorous validation and verification processes are performed to ensure that automated recovery measures do not introduce additional vulnerabilities or compromise the security of patient information [5].



Source: Adapted and combined from conceptual models discussed in [1], [3], and [4]

Figure 1 AI-Based Self-Healing Architecture in Cloud and IoT Environments

7. Bio-Inspired Models and Nature-Inspired Healing Mechanisms

Biological systems have increasingly inspired the self-healing of software systems. Nature has demonstrated a wide range of self-healing capabilities, such as cellular regeneration and immune responses, which have become paradigms for software healing algorithms. AI models can adopt the notion of biological dynamism in a decentralized manner to assess system health, isolate malfunctioning elements, and initiate recovery.

In bio-inspired architectures, fault detection is performed using immune system-like anomaly detection, in which normal system behavior is learned and anomalies are perceived as potential faults. These anomalies trigger corrective actions that resemble immune responses in living organisms. For example, ant colony optimization, neural regeneration, and evolutionary adaptation-based AI models have been developed to enable software systems to reroute data, restructure dependencies, or replace failed modules without requiring manual restarts or explicit commands [6].

These adaptive models strengthen learning by continuously expanding their learning boundaries. Bio-inspired systems can withstand unforeseen fault patterns, unlike conventional rule-based fault detectors that generalize only from historical occurrences. This capability is particularly valuable in dynamic and unpredictable settings, such as real-time financial trading systems or space missions, where systems must self-adapt during operation [6].

The interaction between AI and biological models has expanded the scope of self-healing software. It suggests a transition from fixed architectures to more organic systems with learning capabilities that become increasingly intelligent and autonomous through experience [6].

8. Machine Learning in Log Analysis and Proactive Diagnosis

Logs remain a valid and crucial source of information for debugging software errors, tracking system usage, application health, error messages, and user actions. However, the volume of log data in modern systems makes manual analysis impractical. As a result, machine learning models are used to analyze logs in real time, identify anomalies, and proactively address faults.

The use of AI for log analysis employs natural language processing (NLP), recurrent neural networks (RNNs), and clustering algorithms to detect patterns and deviations. These models sort logs, rank anomalies by severity, and

recommend or initiate recovery actions. One key advantage is the early identification of fault precursors, which helps prevent system degradation [7].

Supervised machine learning models are typically trained on historical log data to identify known fault patterns. Over time, the models can detect previously unseen patterns and classify them based on similarity to known anomalies. In addition, anomaly detection techniques are scalable to slow-developing performance issues, such as memory leaks or disk space saturation, which are difficult to detect using traditional monitoring approaches [7].

A common application scenario is microservice-based systems, in which logs are generated by multiple interacting components. AI models compare logs across services to identify the sources of cascading failures. This reduces the mean time to resolution (MTTR) and enhances overall system reliability. These models are also integrated with alerting systems, enabling real-time notifications to administrators and the automatic activation of corrective workflows [7].

9. Autonomous Quality Assurance in Software Testing

Testing is a significant step in the software development life cycle and is required to enhance reliability, although it is time-consuming and subject to error. Another consequence of AI adoption has been self-regulating quality assurance, whereby test code and test environments can automatically evolve as the system changes. Self-healing in this context is a form of test automation that focuses on the dynamic updating of failing test cases based on changes in the user interface, APIs, or system workflows.

Execution logs and user flows are analyzed automatically by AI-based test automation tools. When a test fails due to a minor system update (for example, an element ID being renamed or a button changing position), the AI model analyzes the failure, identifies the most probable cause, and attempts to adjust the test script to match the new system conditions. This approach helps reduce testing time and ensures that significant defects are detected and corrected even as the system evolves over time [8].

Additionally, test coverage is maximized using reinforcement learning. Over time, the system learns which components are most likely to fail and allocates additional testing resources to them. This targeted testing enhances the effectiveness of fault detection and minimizes unnecessary test repetition. As a result, manual intervention is reduced, enabling software teams to deliver higher-quality releases more efficiently [8].

10. Machine Learning-Driven Recovery and Engineering Transformation

The advent of AI in the software engineering sector has transformed traditional recovery systems from reactive approaches into proactive and predictive ones. Today, machine learning models trained on system telemetry and behavioral data can manage entire recovery operations. These include workload migration, dynamic resource allocation, and automated rollback when deployment failures occur.

Machine learning can be used to enhance system situational awareness by continuously analyzing metrics such as CPU usage, memory utilization, latency, and error rates. When baseline patterns are violated, the AI system evaluates the potential impact and recommends a healing plan with a high probability of restoring normal operation. This approach not only reduces the frequency of service disruptions but also improves resource utilization [9].

These systems are particularly valuable in high-availability configurations where downtime is unacceptable. Previously unpredictable or difficult-to-diagnose faults can be managed more effectively because AI-controlled systems reduce dependence on predefined rule sets and enable more adaptive fault response mechanisms [9].

11. Advancements in Anomaly Detection and Resilience Engineering

Anomaly detection is one of the pillars of self-healing software, as it enables a system to identify irregularities early before they develop into severe malfunctions. Recent advances in deep learning and unsupervised learning models have made significant contributions to improving the accuracy and efficiency of anomaly detection systems.

The AI-based anomaly detection models used in self-healing software are typically streaming models that continuously process performance metrics, logs, and user activities. These models commonly employ autoencoders, long short-term memory (LSTM) networks, and Gaussian mixture models to learn normal patterns and identify deviations. When an anomaly is detected, the system either alerts an operator or automatically initiates a predefined healing process [10].

Graph-based anomaly detection is also gaining popularity, particularly in systems with highly interdependent components. In this approach, the system is represented as a graph, where nodes correspond to services or components and edges represent communication or dependencies. AI models monitor the evolution of these graphs to identify abnormalities, such as sudden disconnections or unexpected traffic spikes [10].

By integrating anomaly detectors into system architectures, organizations can enhance resilience and reduce operational risks. Furthermore, adaptive learning mechanisms allow models to improve over time, enabling them to accommodate system updates and changes in user behavior [10].

12. Semantic AI for Cognitive Healing in Cloud Infrastructure

Semantic AI introduces a situational and interpretative dimension to fault diagnosis and self-repair. In cognitive healing, semantic AI is capable of analyzing not only quantitative metrics and logs but also the meaning and purpose of system interactions. This approach enables systems to understand failures from a broader perspective, detect both minor and critical faults, and select the most appropriate recovery plans.

Knowledge within such systems is represented using knowledge graphs, ontologies, and semantic networks, which support cognitive self-healing. The AI engine leverages these knowledge structures to encode problems and determine appropriate actions in the event of errors. For example, in a cloud environment, an increase in latency may have different implications depending on workload, time of day, and user priority. Semantic AI considers these contextual factors, leading to more accurate and context-specific recovery decisions [11].

Such systems are valuable in fault management because they reduce false positives and enhance interpretability. They also contribute to the development of intelligent service orchestration engines capable of reasoning about dependencies and proactively allocating resources [11].

13. Self-Healing in the Modern Software Development Lifecycle

Continuous integration and continuous deployment (CI/CD) models demand rapid and reliable software releases. To ensure a stable, yet fast, development lifecycle, AI-driven capabilities are being integrated throughout the software development process. AI-based testing tools detect flaky tests and automatically repair or isolate them, while suspicious builds are prevented from being promoted to production from staging environments.

Software environments are becoming increasingly complex, with self-healing tools integrated into development platforms and DevOps toolchains. These tools monitor system conditions, collect telemetry data, and update predictive models in real time. This continuous learning enables systems to adapt to code changes, the introduction of new features, and configuration updates, helping to maintain system health across software versions [12].

In addition, developers can use AI models to assist in proposing bug fixes, code changes, and configuration updates that help prevent common errors. This makes the software development process more robust and effective, allowing faults to be identified and resolved at earlier stages of the development lifecycle [12].

Table 1 Key AI Techniques Used in Self-Healing Software Systems

AI Technique	Application in Self-Healing Systems	Benefits
Supervised Learning	Fault classification using labeled data	High accuracy in known fault types
Unsupervised Learning	Anomaly detection in unlabeled data	Identification of unknown fault patterns
Reinforcement Learning	Dynamic policy learning for recovery decisions	Continuous improvement in healing strategies
Natural Language Processing	Log analysis and root cause identification	Improved fault diagnosis from unstructured data
Semantic AI	Context-aware fault prediction and cognitive healing	Intelligent decision-making with contextual understanding

Source: Synthesized from [1] through [12]

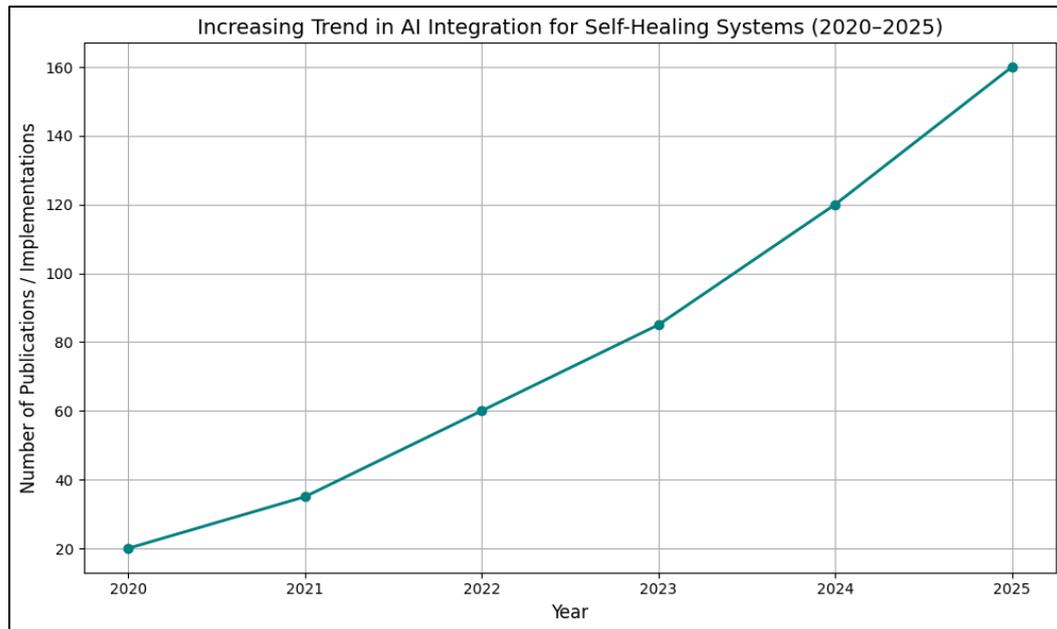


Figure 2 Increasing Trend in AI Integration for Self-Healing Systems (2020–2025)

Source: Compiled from data and trend analysis across [1] to [12]

A line graph depicting the growing number of publications and industry implementations of AI-driven self-healing systems between 2020 and 2025, based on cumulative studies from the referenced sources.

14. Conclusion

The combination of AI and software engineering has introduced a new era of autonomous and resilient systems with capabilities for self-monitoring, fault prediction, and automatic repair. A major paradigm shift in system reliability is the transition from reactive fault handling to proactive and context-aware healing. AI-based self-healing systems are transforming domains such as cloud computing, IoT, and healthcare through the use of machine learning, semantic reasoning, bio-inspired techniques, and real-time analytics.

Key challenges that require continued attention include model interpretability, data privacy, and the validation of autonomous actions, particularly as these systems continue to evolve. Increased adoption of AI in software systems promises greater resilience, reduced downtime, and improved efficiency in an increasingly digital and interconnected world.

References

- [1] Kambala, G. (2025, May). AI-Driven Fault Detection and Self-Healing Framework for Resilient Distributed Software Systems in Mission-Critical Application. In *2025 Global Conference in Emerging Technology (GINOTECH)* (pp. 1-7). IEEE.
- [2] Rana, S. (2025). AI-driven fault detection and predictive maintenance in electrical power systems: A systematic review of data-driven approaches, digital twins, and self-healing grids. *American Journal of Advanced Technology and Engineering Solutions*, 1(01), 258-289.
- [3] Jain, M. (2025). Self-Healing Networks with AI-Based Fault Prediction in IoT Ecosystems.
- [4] Anny, D. (2025). Self-Healing Cloud Systems: Integrating AI for Automated Recovery.
- [5] Baloch, J., & Rahman, N. Smart Recovery: Self-Healing Cloud Systems and AI Monitoring in Critical Healthcare Applications.
- [6] Baqar, M., Khanda, R., & Naqvi, S. (2025). Self-Healing Software Systems: Lessons from Nature, Powered by AI. *arXiv preprint arXiv:2504.20093*.

- [7] Ali, T., Iqbal, R., Ansari, N. M., Tariq, T., & Rafique, A. A. (2025). AI-POWERED ANOMALY DETECTION IN SOFTWARE LOGS: A MACHINE LEARNING APPROACH FOR PROACTIVE FAULT DIAGNOSIS AND SELF-HEALING SYSTEMS. *Spectrum of Engineering Sciences*, 3(3), 302-322.
- [8] Abhichandani, S., Vadrevu, N. R. T., & Bagmar, V. (2025, June). AI-Driven Self-Healing in Test Automation: A Review of Autonomous Quality Assurance. In *2025 3rd International Conference on Inventive Computing and Informatics (ICICI)* (pp. 1601-1608). IEEE.
- [9] Shah, H., & Patel, J. Machine Learning-Driven Self-Healing Systems: Revolutionizing Software Engineering.
- [10] Loveth, R. G. AI-Driven Anomaly Detection: A New Era In Self-Healing Software Engineering.
- [11] Anand, S., Miglani, S., & Anand, R. (2025). AI-Driven Cognitive Fault Detection and Self-Healing in Cloud Infrastructure Using Semantic Understanding. In *Establishing AI-Specific Cloud Computing Infrastructure* (pp. 319-330). IGI Global Scientific Publishing.
- [12] Mohapatra, P. (2025). Artificial Intelligence in Test Automation and Self-Healing. *Intelligent Assurance: Artificial Intelligence-Powered Software Testing in the Modern Development Lifecycle*, 4, 93.