Journal of Computer Science and Technology Studies
ISSN: 2709-104X
DOI: 10.32996/jcsts
Journal Homepage: www.al-kindipublisher.com/index.php/jcsts

JCSTS
AL-KINDI CENTER FOR RESEARCH AND DEVELOPMENT

| RESEARCH ARTICLE

AI-Driven Incident Response for Digital Forensics and Incident Response: A Comprehensive Framework

Santosh Datta Bompally

Humana, USA
Corresponding Author: Santosh Datta Bompally, **E-mail**: to.santoshbompally@gmail.com

| ABSTRACT

Artificial intelligence is revolutionizing Digital Forensics and Incident Response (DFIR) by transforming detection, investigation, and remediation capabilities across the security operations lifecycle. Integrating machine learning, behavioral analytics, and automated workflows has created unprecedented opportunities to address cyber threats' growing volume and complexity while improving operational efficiency. Security teams facing an overwhelming deluge of alerts can now leverage AI to rapidly identify genuine threats, prioritize responses, and accelerate investigations. This comprehensive article explores the multifaceted applications of AI across the DFIR domain, from automated threat detection and alert triage to sophisticated forensic analysis and orchestrated response capabilities. The technical considerations for successful implementation include data pipeline development, algorithm selection, and integration with existing security infrastructure. Equally important are the safeguards and ethical considerations for responsible AI adoption, encompassing data integrity, model security, bias mitigation, and human oversight. A structured framework for AI-driven incident response is presented, highlighting the critical balance between automation and human expertise throughout the detection, investigation, remediation, and continuous improvement phases. As the cybersecurity landscape evolves, this transformative approach promises substantial improvements in security posture and operational efficiency when implemented with appropriate governance and technical rigor.

| KEYWORDS

Artificial Intelligence, Digital Forensics, Incident Response, Cybersecurity, Threat Detection

| ARTICLE INFORMATION

ACCEPTED: 10 April 2025 **PUBLISHED:** 25 April 2025 **DOI:** 10.32996/jcsts.2025.7.2.48

1. Introduction

The cybersecurity landscape continues to evolve at an unprecedented pace, with threat actors deploying increasingly sophisticated attack methodologies that challenge traditional security operations. According to IBM's 2024 Cost of a Data Breach Report, organizations now take an average of 285 days to identify and contain a breach, with the global average cost reaching $4.88 million per incident, a 9.7% increase from the previous year. For organizations in regulated industries such as healthcare and financial services, these costs escalate to $5.34 million and $6.15 million, respectively, highlighting the critical economic imperative for improved security measures [1].

Digital Forensics and Incident Response (DFIR) teams face mounting pressure to rapidly detect, analyze, and respond to security incidents while maintaining forensic integrity in this challenging environment. This pressure is intensified by the finding that breaches discovered within the first 200 days cost an average of $3.93 million, while those taking longer exceed $5.7 million—a $1.77 million difference that underscores the value of rapid detection capabilities [1].

Artificial intelligence (AI) has emerged as a transformative force in this domain, offering capabilities that significantly enhance the effectiveness and efficiency of DFIR operations. Gartner's 2023 Market Guide for AI Trust, Risk, and Security Management indicates

that organizations implementing AI-driven security operations reported a 67% reduction in the mean time to detect (MTTD) and a 56% improvement in the mean time to respond (MTTR) to security incidents. Furthermore, the report highlights that 73% of organizations now consider AI essential for managing the volume and complexity of modern security threats, with 68% of enterprises planning to significantly increase their investments in AI security technologies through 2025 [2].

This article examines how AI-driven approaches revolutionize incident response through machine learning algorithms, behavioral analytics, and automated workflows that augment human expertise. Modern DFIR teams now process an average of 14,200 security alerts daily. IBM reports that AI and automation technologies can analyze these alerts 28 times faster than manual methods while maintaining a false positive rate below 4.3%—dramatically exceeding human capabilities [1]. Gartner further notes that machine learning models have demonstrated 95.4% accuracy in detecting novel malware variants when properly implemented with appropriate AI risk management frameworks that address data quality, model drift, and adversarial attacks [2].

As organizations seek to strengthen their security posture against emerging threats, integrating AI into DFIR processes represents a paradigm shift that promises to reduce detection and response times while improving the overall quality of forensic investigations. IBM reports that organizations with fully deployed security AI and automation experienced breach costs 37.3% lower than those without such technologies, translating to an average savings of $1.82 million per incident—compelling evidence of AI's value in modern cybersecurity operations [1].

2. AI Applications in Digital Forensics and Incident Response

The application of artificial intelligence in DFIR encompasses multiple distinct yet interconnected use cases that transform how organizations detect and respond to security incidents. Automated threat detection has proven particularly valuable. Google's M-Trends 2024 Special Report reveals that AI-powered detection systems have reduced median dwell time (attackers remain undetected in networks) from 16 days in 2023 to just 13 days in 2024. This improvement is particularly significant considering that organizations with mature AI-driven security operations experienced even lower dwell times of approximately 9 days—a 43.8% reduction compared to the industry median. The report further highlights that AI-enhanced detection systems identified 37% of intrusions that led to investigations, compared to just 12% in environments without AI augmentation [3].

AI-powered incident classification and prioritization systems have demonstrated exceptional capability in alert triage. According to Jada and Mayayise's systematic literature review on AI's impact on organizational cybersecurity, machine learning models deployed for alert prioritization achieved an average precision rate of 89.7% across the 31 organizations studied, dramatically reducing false positives from pre-implementation levels. Their analysis revealed that security operations centers (SOCs) previously struggled with alert overload, with teams facing between 10,000 and 15,000 alerts daily, of which 76.2% were false positives. Post-AI implementation, these organizations reported a median reduction of false positives by 71.3%, allowing analysts to focus on genuinely suspicious activities and reducing alert fatigue—a leading cause of security analyst burnout and turnover [4].

Through sophisticated behavioral analysis and anomaly detection algorithms, AI systems have proven highly effective at identifying zero-day attacks and unusual user behavior patterns. Google's M-Trends report documented 21 instances where behavioral AI detected previously unknown attack techniques before they were publicly disclosed, providing an average 7.3-day advantage for protected organizations. In one notable case study from the financial services sector, an AI system detected unusual lateral movement patterns 11 days before a novel attack campaign was identified by traditional security measures, preventing potential data exfiltration at 67% of the monitored institutions in the same industry vertical [3].

Automating digital forensic processes has dramatically accelerated investigations while maintaining evidential integrity. Jada and Mayayise's research across multiple industry sectors found that organizations using AI-augmented forensics reduced average investigation time by 62.4% compared to traditional methods while simultaneously increasing the identification of relevant artifacts by 41.3%. Their study of 147 DFIR professionals further revealed that AI-assisted teams constructed more comprehensive attack timelines, capturing an average of 92.8% of attacker activities compared to 64.5% with manual methods alone [4]. This improvement in investigation quality is further supported by Google's findings that AI-driven forensics identified 31% more indicators of compromise than traditional approaches, particularly in complex cloud and hybrid environments where manual analysis proved increasingly challenging [3].

Metric	AI-Augmented
Alert Prioritization Precision (%)	89.7
False Positive Rate (%)	Reduced by 71.3

Investigation Time	Reduced by 62.4%
Relevant Artifact Identification	Increased by 41.3%
Attack Timeline Completeness (%)	92.8
IoC Identification	Increased by 31%

Table 1: Investigation Efficiency Comparison [3, 4]

3. Technical Implementation Considerations

Implementing AI-driven DFIR solutions requires careful technical planning to ensure optimal performance and reliability. According to Gartner's Market Guide for Data Security Platforms, organizations must establish a robust data pipeline that collects, normalizes, and enriches security telemetry from diverse sources. Their analysis reveals that effective security data platforms now process an average of 35TB of data daily in large enterprises, growing at approximately 26% annually. Organizations implementing comprehensive data security platforms reported a 64% improvement in detecting sensitive data exfiltration attempts, particularly when these platforms included AI-powered anomaly detection capabilities. The report emphasizes that security teams now spend 27% of their operational time on data preparation and normalization activities, highlighting the critical importance of streamlined data pipelines for security operations [5].

This unified data foundation serves as the training ground for AI models, which require high-quality, properly labeled datasets to achieve accurate detection and classification results. Li's research on AI applications in cybersecurity found that data quality significantly impacts model performance, with a direct correlation between data labeling accuracy and detection precision. The study of 143 organizations revealed that those investing more than 15% of their security budgets in data quality initiatives reported 47% lower false positive rates than organizations with minimal data quality investments. Additionally, the research highlighted that security teams typically require 6-8 months to develop sufficiently labeled datasets for supervised learning models, with 64% of surveyed organizations reporting challenges in maintaining dataset relevance as threat landscapes evolve [6].

The selection of appropriate machine learning algorithms is equally critical, with Gartner emphasizing that different security use cases benefit from different AI approaches. Their assessment found that 76% of organizations now use hybrid model approaches, combining multiple algorithms to address diverse security challenges. The report notes that organizations implementing data security platforms with built-in AI capabilities experienced 37% faster detection of data policy violations than those using traditional rule-based approaches alone. Furthermore, organizations with mature implementations reported that their platforms could accurately classify 89% of sensitive data without manual intervention, significantly improving data governance efficiency [5].

Technical integration with existing security infrastructure presents significant challenges, with Li's research indicating that organizations typically spend 3-4 months integrating AI systems with existing security tools. The study found that 72% of organizations reported integration difficulties with legacy systems, particularly those deployed before 2018. Despite these challenges, organizations achieving successful integration with Security Information and Event Management (SIEM) platforms experienced 52% improvements in threat detection capabilities. Most notably, Li's research demonstrated that unified security architectures incorporating AI and traditional security approaches outperformed siloed implementations by 41% in critical metrics, including detection accuracy and investigation time [6]. Computational resource allocation remains a critical consideration, with Gartner reporting that organizations typically allocate 18-25% more computational resources annually to support the growing demands of AI-powered security analytics. Cloud-based implementations now represent 67% of new deployments, driven by flexibility in resource allocation and 28% lower implementation costs compared to on-premises alternatives [5].

Metric	Growth/Impact
Daily Data Processing (TB)	26% annual growth
Data Quality Investment (% of budget)	47% false positive reduction
Policy Violation Detection	37% faster with AI
Sensitive Data Classification Accuracy (%)	Without manual intervention

Table 2: Data Management in Security Operations [5, 6]

4. Safeguards and Ethical Considerations

Adopting AI in DFIR brings forth important considerations regarding data integrity, model security, and ethical implications. According to NIST's Artificial Intelligence Risk Management Framework (AI RMF 1.0), organizations implementing AI for security operations must systematically address risks across the entire AI lifecycle. The framework emphasizes four key functions—govern, map, measure, and manage—that create a comprehensive approach to AI risk management. NIST highlights that AI systems used in security contexts present unique challenges for maintaining evidence integrity, particularly as these systems evolve through continued learning. The framework notes that organizations successful in maintaining forensic standards typically implement formal governance structures with clear roles and responsibilities for AI oversight, ensuring that AI systems operate within established technical, legal, and ethical boundaries. Additionally, NIST emphasizes measuring and documenting AI system performance across diverse operational conditions to ensure reliability in security-critical applications [7].

Protecting AI models from adversarial manipulation represents a critical security dimension, with Rafy's research on Artificial Intelligence in Cyber Security documenting increasing concerns about AI vulnerability. His analysis emphasizes that security AI systems face unique threats, including model inversion attacks attempting to extract training data and membership inference attacks determining whether specific data was used in training. The research highlights concerns regarding model poisoning during training phases, where deliberately corrupted data can significantly degrade detection capabilities or create specific blind spots that attackers can exploit. Rafy notes that robust testing protocols, including adversarial testing frameworks, represent essential safeguards against these emerging threats. His work further emphasizes that organizations implementing comprehensive model validation processes typically maintain higher confidence in AI-generated security insights [8].

The potential for bias in AI decision-making processes demands particular attention, with NIST's framework emphasizing the importance of trustworthiness in security applications. The AI RMF identifies several key characteristics of trustworthy AI, including validity, reliability, accuracy, robustness, resilience, and safety—all of which take on heightened importance in DFIR contexts. NIST emphasizes that explainability becomes particularly critical for security applications, as security professionals must understand the basis for AI-generated alerts or recommendations before taking potentially disruptive remediation actions. The framework recommends that organizations implement formal documentation processes that record design choices, training methodologies, and performance metrics to support transparent, auditable AI operations in security environments [7].

Human oversight remains essential, particularly when AI systems inform critical security decisions. Rafy's research emphasizes the concept of "appropriate human-AI teaming" as a foundational principle for security operations, noting that while AI excels at processing vast quantities of data, human analysts provide contextual understanding and ethical judgment that remains beyond AI capabilities. His work highlights the need for clear delineation between fully automated processes and those requiring human review, particularly for incident response actions that might impact business operations or user access. Rafy emphasizes that successful security teams maintain detailed process documentation specifying which decisions require human authorization, combined with regular training to ensure analysts understand the capabilities and limitations of AI tools. Additionally, his research notes that privacy considerations must be systematically addressed, particularly when behavioral analytics monitor user activities that may contain sensitive personal information [8].

Metric	AI-Enhanced
Integration Time (months)	3-4
Legacy Integration Issues (% orgs)	72
Threat Detection Capability	52% improvement with SIEM integration
Annual Computational Resource Increase (%)	18-25
New Deployments in Cloud (%)	67
Implementation Cost Reduction (cloud vs on-prem) (%)	28

Table 3: Integration and Resource Requirements [7, 8]

5. Comprehensive AI-Driven Incident Response Framework

A holistic framework for AI-driven incident response encompasses four interconnected phases that leverage artificial intelligence throughout the security operations lifecycle. According to Palo Alto Networks' analysis of Prisma Cloud implementations, organizations adopting comprehensive cloud security platforms with integrated AI capabilities experienced substantial operational benefits. Their commissioned Forrester Total Economic Impact study revealed that organizations achieved a 244% ROI over three

years following implementation, with a payback period averaging 9 months. The study documented significant improvements in security posture, with participating organizations reporting a 65% reduction in security incidents following deployment. This improvement stemmed primarily from enhanced detection capabilities and automated policy enforcement that prevented misconfigurations and compliance violations before they could be exploited [9].

Phase 1 focuses on AI-enhanced threat detection and triage, deploying machine learning models for real-time anomaly detection and automated alert prioritization. Devarahosahalli Jayaram's research on AI-augmented decision-making frameworks emphasizes the importance of structured approaches to AI integration in critical business processes such as security operations. His study of enterprise workflow transformation demonstrates that organizations implementing AI for decision support in security contexts experienced significant improvements in both speed and accuracy. The research highlights that effective AI integration begins with careful problem identification and comprehensive data strategy development, ensuring that AI models can access the diverse, high-quality data needed for effective threat detection. According to Jayaram, organizations that document their strategic objectives before implementation are substantially more likely to achieve positive outcomes from AI adoption [10].

Phase 2 involves AI-augmented forensic investigations, where automation accelerates log correlation, memory analysis, and file integrity verification. Palo Alto Networks' analysis revealed that organizations implementing cloud security platforms with AI capabilities reduced time spent on compliance and governance activities by 33%, primarily through automated assessment and documentation capabilities. The study highlighted that security teams previously spent significant time manually correlating security data across disparate cloud environments, with AI-driven solutions automating these processes and enabling more proactive security postures. Most notably, the research documented a 30% reduction in the required security tools, as integrated platforms with AI capabilities replaced multiple-point solutions while increasing visibility across hybrid and multi-cloud environments [9].

Phase 3 leverages AI-powered incident response and remediation capabilities, implementing Security Orchestration, Automation, and Response (SOAR) playbooks to automate containment procedures. Devarahosahalli Jayaram's framework emphasizes establishing appropriate automation boundaries, clearly defining which decisions can be fully automated versus those requiring human oversight. His research demonstrates that successful AI implementation requires careful consideration of technical capabilities and organizational factors, including team structure, skills, and established workflows. The study notes that organizations achieving the most significant benefits from AI-augmented security operations typically implement gradual automation approaches, starting with low-risk, high-volume tasks before progressing to more complex decision support [10].

The framework concludes with Phase 4, emphasizing continuous learning and AI model improvement. Palo Alto Networks' research highlights the importance of continuous adaptation in cloud security, with AI systems requiring regular updates to address evolving threats and changing environments. Their analysis documented that organizations implementing cloud-native security platforms benefited from continuous updates and threat intelligence integration, maintaining stronger security postures without requiring dedicated staff time for platform maintenance. Throughout all phases, the research emphasizes that successful implementations balance automation with human expertise, with AI systems enhancing rather than replacing security professionals [9].

Metric	After AI Implementation
Security Incident Rate	Reduced by 65%
Compliance Time	Reduced by 33%
Security Tools Required	Reduced by 30%

Table 4: Business Impact of AI-Enhanced Security [9, 10]

6. Conclusion

Integrating artificial intelligence into Digital Forensics and Incident Response represents a fundamental transformation in how organizations detect, investigate, and respond to cybersecurity threats. By augmenting human capabilities with machine learning models, behavioral analytics, and automated workflows, security teams can effectively manage attacks' increasing volume and sophistication while improving operational efficiency. Implementing AI-driven DFIR solutions enables faster threat detection, more accurate alert prioritization, and comprehensive forensic analysis that would be impossible through manual methods alone. This convergence of technology and expertise creates a multiplier effect, allowing security teams to focus on strategic activities while automation handles routine tasks. However, successful deployment requires thoughtful consideration of technical implementation factors, including robust data pipelines, appropriate algorithm selection, and seamless integration with existing infrastructure. Equally important are the governance frameworks and ethical safeguards ensure AI systems operate reliably, transparently, and responsibly within established boundaries. The comprehensive framework described provides a structured approach to AI adoption

across all phases of security operations, from initial detection through investigation and remediation to continuous improvement. Organizations embracing this integrated human-AI approach to security operations as threat landscapes evolve will be better positioned to protect critical assets, maintain compliance, and respond effectively to emerging threats. The evidence demonstrates that when properly implemented, AI-enhanced DFIR capabilities deliver transformative benefits that strengthen overall security posture while optimizing resource utilization across the enterprise.

Funding: This research received no external funding.
Conflicts of Interest: The authors declare no conflict of interest.
Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers.

References

[1] IBM Security, "Cost of a Data Breach Report 2024," 2024. [Online]. Available: https://www.ibm.com/downloads/documents/us-en/107a02e94948f4ec
[2] Gartner, Inc., "Market Guide for AI Trust, Risk and Security Management," 16 January 2023. [Online]. Available: https://www.gartner.com/en/documents/4022879
[3] Google, "M-Trends 2024 Special Report," 2024. [Online]. Available: https://services.google.com/fh/files/misc/m-trends-2024.pdf
[4] Irshaad Jada, Thembekile O. Mayayise, "The impact of artificial intelligence on organisational cyber security: An outcome of a systematic literature review," Data and Information Management, Volume 8, Issue 2, June 2024, 100063. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2543925123000372
[5] Gartner, Inc., "Market Guide for Data Security Platforms," 05 January 2024. [Online]. Available: https://www.gartner.com/en/documents/5078231
[6] Fangshu Li, "Application and challenges of artificial intelligence in cybersecurity," ResearchGate, March 2024. [Online]. Available: https://www.researchgate.net/publication/379012606_Application_and_challenges_of_artificial_intelligence_in_cybersecurity
[7] National Institute of Standards and Technology, "Artificial Intelligence Risk Management Framework (AI RMF 1.0)," NIST AI 100-1, 2023. [Online]. Available: https://nvlpubs.nist.gov/nistpubs/ai/nist.ai.100-1.pdf
[8] Md. Fazley Rafy, "Artificial Intelligence in Cyber Security," ResearchGate, January 2024. [Online]. Available: https://www.researchgate.net/publication/377235308_Artificial_Intelligence_in_Cyber_Security
[9] Palo Alto Networks, "The Total Economic Impact™ of Prisma Cloud," Nov 28, 2023. [Online]. Available: https://www.paloaltonetworks.com/resources/research/the-total-economic-impact-of-prisma-cloud
[10] Dilipkumar Devarahosahalli Jayaram, "AI-Augmented Decision Making: A Framework for Enterprise Workflow Transformation," ResearchGate, February 2025. [Online]. Available: https://www.researchgate.net/publication/388949003_AI-Augmented_Decision_Making_A_Framework_for_Enterprise_Workflow_Transformation