The American Journal of Engineering and Technology
ISSN 2689-0984 Volume 08 - 2026

Adaptive Trust: A Comparative Analysis of Cybersecurity Metrics and AI-Driven Privacy Safety Enforcement. Traditional Fidelity versus AI-Driven Velocity

Savi Grover
Independent Researcher Rahway, New Jersey, USA

Received: 24 Jan 2026 | Received Revised Version: 13 Feb 2026 | Accepted: 18 Mar 2026 | Published: 23 Apr 2026

Volume 08 Issue 04 2026 | Crossref DOI: 10.37547/tajet/Volume08Issue04-13

Abstract

Cybersecurity Testing and Evaluation (T&E) comprise of a foundational resilience component, moving beyond simple quality assurance to become a critical process for continuous organizational hardening. Effective T&E enhances emergency plans, policies, attacks resistance, filtration, firewall strengthening procedures, promoting the efficient utilization of capabilities required to respond to sophisticated cyber-attacks. In this paper, we are performing comparative analysis of traditional security and cyber evaluation metrics with upcoming AI driven enhanced secure measures. AI-LLM security techniques like early defect prediction, defect clustering, Secure cloud and Automated Incident Response measures are weighted against traditional security techniques in terms of – speed, velocity, criticality and depth of coverage scope. These metrics point towards countless advantages of combining the two for greater holistic impact.

Keywords: Cybersecurity Testing, T&E, AI Safety, AI-LLM Security Methodologies

Cite This Article: Savi Grover. (2026). Adaptive Trust: A Comparative Analysis of Cybersecurity Metrics and AI-Driven Privacy Safety Enforcement. Traditional Fidelity versus AI-Driven Velocity. The American Journal of Engineering and Technology, 8(4), 128–135. https://doi.org/10.37547/tajet/Volume08Issue04-13

1. Introduction

The National Institute of Standards and Technology (NIST) methodologies define T&E events, distinguishing between Tests—which use quantifiable metrics to validate the operability of an IT system or component in a specified operational environment—and Exercises, which are broader events designed to validate planning elements in the context of cyber incidents [1]. As these digital adversaries evolved, so did the arsenal of cybersecurity defenders. Antivirus software, firewalls, and intrusion detection systems became essential tools for network defense [2].

The necessity for continuous, highly efficient T&E methodologies is underestimated by the rapidly accelerating threat calculus—an emerging era where Artificial Intelligence (AI) fights AI and AI enhances AI. Organizations are shifting into a turning point characterized by autonomous, self-optimizing AI threat actors capable of planning, executing, and refining sophisticated campaigns with minimal human oversight [3]. This technological evolution places immense strain on traditional, periodic T&E models. If a typical penetration test occurs quarterly or even annually [3], the window of vulnerability remains wide open for months. For example- With healthcare data breaches costing an

average of $9.77 million per incident in 2024, the highest across all industries for the 14th consecutive year, the need for robust security is paramount [4]. Given the unprecedented speed at which AI can automate attack execution and exploitation workflows [5], the time delay between a vulnerability's emergence and its potential exploitation shrinks drastically. Consequently, any T&E methodology that is not near real-time or continuous is functionally incompatible with the modern pace of risk, rendering it baseless for the security validation of dynamic large-scale systems. This structural failure of traditional cadence dictates a compulsory shift toward automated, frequent, or continuous validation protocols, regardless of initial investment costs [6].

Current Problem Statement- A robust T&E program must be defined by a comprehensive risk assessment methodology [7]. This methodology establishes an explicit risk model, defines key assessable risk factors and their interrelationships, and specifies the assessment approach, which may be quantitative, qualitative, or a combination of both. Furthermore, the analysis approach, such as threat-oriented, asset/impact-oriented, or vulnerability-oriented, must be determined to ensure adequate coverage of the problem space [7]. Strategic decisions regarding T&E are fundamentally constrained by necessary trade-offs (TTA) in security engineering. Cybersecurity professionals constantly balance the need for security improvements against other critical factors, including usability, cost, privacy, openness, convenience and system performance. For instance, implementing stringent security measures often requires significant investment, impacting budgets, or may compromise end-to-end system performance, which can result in client loss in (B2C) services [8].

Paper Objectives: This comparative analysis evaluates traditional T&E methodologies across four distinct and essential dimensions in comparison to modern AI-ML safety trends and security metrics.

1. **Velocity (Speed & Cadence):** How quickly and how frequently testing can be performed and how fast results are delivered.
2. **Coverage (Breadth):** The expanse of the digital estate and attack surface analyzed.
3. **Depth (Context & Creativity):** The ability to find unknown, complex business logic flaws, custom protocol vulnerabilities, and chained exploits that require creative insight.
4. **Fidelity (Safety & Trust):** The trustworthiness and accuracy of the findings, principally measured by the false positive rate (FPR).

2. Traditional Foundation T&E Methods: The Quality and Trust Baseline

2.1. Automated Scanning: Static (SAST) and Dynamic (DAST) Analysis

Traditional automated T&E primarily relies on Static Application Security Testing (SAST) and Dynamic Application Security Testing (DAST). SAST provides quick and efficient results by examining the malware or application's code and structure *without* executing it. SAST detects issues related to coding standards violations, potential security vulnerabilities, and logical errors in code structure, proving effective at finding clear, traceable issues such as injection flaws or buffer overflow vulnerabilities [9]. Conversely, DAST provides in-depth insights by observing the code running in a controlled environment (runtime analysis), detecting problems that only occur during execution, such as memory leaks and performance bottlenecks [10]. ML based automations like- two approaches are used: Code Metrics and Feature Extraction. The amalgamation of these two approaches helps us to identify and analyze the software. The approach was tested upon 100 releases and out of these, 80 releases were found with vulnerable code in it [11].

However, traditional automated scanners are hindered by significant limitations, particularly concerning Fidelity. These tools depend on signatures and rules and often produce high false-positive rates. This inaccuracy stems from the tool's inability to verify the data flow or safety within the application, particularly when external systems or closed-source components are involved [12]. Without access to the underlying code of these components, the tool cannot definitively determine the security posture throughout the data lifecycle, frequently resulting in false positives. This systemic inefficiency creates a substantial operational cost, as development teams waste hours trying to remediate non-existent issues [13]. Another drawback is the inability to grasp complex context. Traditional automated analysis tools are less effective at identifying intricate logic errors or complex bugs [12]. The tools may achieve broad coverage and high speed, but this efficiency is often superficial, failing to understand how vulnerabilities interconnect or what is genuinely exploitable. This established friction points in

the traditional security lifecycle, driven by poor data quality, created a compelling market necessity for improved automation, shifting the focus of innovation from mere speed to the enhancement of accuracy and Fidelity.

2.2. Manual and Periodic Penetration Testing (MPT)

Manual penetration testing represents the highest standard for Depth and Fidelity in traditional T&E. It is human-driven, creative, and adaptive [14]. A skilled tester brings essential creativity, intuition, and contextual awareness that no automated tool can currently replicate, excelling at finding unknown vulnerabilities, subtle flaws in business logic, and complex chained exploits that automated scanners are fundamentally blind to. Vulnerability Assessment and Penetration Testing (VAPT) techniques help them to go looking out security loopholes. These security loopholes could also be utilized by attackers to launch attacks on technical assets [15]. MPT provides the gold standard for trust due to its focus on Fidelity. Every finding in a manual report is validated and exploited by a human, delivering virtually zero false positives [16]. This saves significant time and effort for development teams by eliminating the need to chase non-existent issues.

Despite its superior Fidelity and Depth, MPT is severely constrained by Velocity and Coverage. Manual testing is inherently slower and more expensive, limited by the time and skill of the individual tester or team [16]. New vulnerabilities arise constantly, evolving menace at new levels, meaning the security posture degrades immediately after a test is completed, leaving risks exposed for extended periods. These constraints prevent MPT from scaling effectively to cover the vast and dynamically changing attack surfaces of modern enterprise environments. Moreover, the implementation of AI in testing frameworks can significantly reduce the burden of manual testing, allowing teams to allocate resources towards more strategic initiatives that drive innovation and quality [17].

2.3 Formal Methods and Verification (FV): Deterministic Safety

For systems demanding the absolute highest level of assurance, such as mission-critical software or sensitive financial applications, Formal Methods and Verification (FV) serve as the ultimate benchmark for safety. FV employs rigorous mathematical proof techniques to demonstrate the trustworthiness and security of critical components [18]. The safety superiority of FV is derived from its deterministic nature: it can theoretically provide a **100% security** for the absence of entire classes of software bugs and security issues. By producing rigorous arguments about security properties, FV dramatically reduces the long-term cost of software maintenance and facilitates compliance with the highest assurance levels of industrial certification [19]. This establishes a critical metric for comparison: deterministic safety, achieved through mathematical proof, is the functional ceiling of trustworthiness. Since all data-driven AI systems are probabilistic and based on learned correlation, they cannot achieve this level of deterministic assurance. Therefore, while AI may be demonstrably more accurate *on average* at scale, it can never achieve the deterministic security guaranteed by formal verification.

3. AI, LLM, and Agentic Influence on T&E Measures

3.1. AI-Driven Vulnerability Discovery and Penetration Testing

The integration of Artificial Intelligence, specifically Large Language Models (LLMs) and advanced Machine Learning (ML), is fundamentally redefining the Velocity and Coverage dimensions of T&E. AI-driven systems leverage pattern recognition, analysis of massive datasets (such as network traffic and code), and anomaly detection to continuously adapt and identify unusual behavior that could indicate a zero-day exploit or other unknown threats [20]. Over the past years, software fault prediction has been a topic of increasing interest because of the ever-growing need to identify defects during early development stages in many agile-like development environments. By studying, reviewing and predicting common security attacks, linear regression modal can help find security vulnerabilities. The most common approach to fault prediction on far has been based on static machine learning approaches [21]. Other recent works focused on deep learning and neural networks during fault prediction; nevertheless, their latent computational cost and the requirement of large volumes of labeled data render them less applicable to mid-size Agile groups with constrained resources [21]. AI models identified 85% of hidden zero-day threats in an average of 12 minutes, compared to human analysts who achieved 60% accuracy over 4 hours [20].

3.2. AI enhances the Fidelity of Automated Testing

Furthermore, AI enhances the Fidelity of automated

testing. Traditional automated scanners are plagued by high false positives; however, AI-driven T&E reduced the false positive rate by 30% compared to human efforts in the zero-day study cited [20]. This improved accuracy is achieved through sophisticated algorithms that are learned from historical security data, allowing for more precise identification of vulnerabilities [24]. For offensive security, LLMs streamline the penetration testing workflow, automating the identification and execution of rooting techniques and generating scripts for exploitation [22]. This capability transforms the security team's operational model, shifting their focus from time-consuming manual data gathering to strategic data analysis and rapid remediation, thereby delivering superior overall return on investment (ROI) and greater throughput of actionable insights [22].

3.3. The Safety Paradox: Limitations and Adversarial Robustness Gaps

Despite the significant advantages in Velocity and operational accuracy, AI-driven T&E is subject to inherent limitations concerning Depth and inherent safety concerns. Research indicates that LLMs often operate at a "shallow level" when processing complex code, performing comparably to classifiers trained solely on classic code metrics. This dependency suggests LLMs rely heavily on statistical correlation derived from training data rather than true causal logic or creative anomaly detection, limiting their ability to grasp complex patterns and fully realize their potential in finding highly novel vulnerabilities [23]. Consequently, even state-of-the-art models currently fall short of reliably performing end-to-end penetration testing without harnessing essential human guidance, intuition and oversight.

The most profound challenge to AI's trustworthiness is its Adversarial Fragility. LLMs are highly susceptible to carefully crafted adversarial inputs, such as minor perturbations or rephrased prompts, which can lead to degraded performance and poor generalization to inputs outside their training distribution. Testing resilience against these inputs is crucial, often requiring specialized benchmarks [24]. Furthermore, the ethical landscape of AI-driven T&E is complex. LLMs pose risks of misinformation generation and potential privacy violations by inadvertently leaking Personally Identifiable Information (PII) extracted from training data, enabling digital impersonation and financial fraud [26]. The observed structural limitation of LLMs relying on correlation confirms that they will struggle to discover zero-day vulnerabilities that fundamentally break existing logic in unforeseen ways.

3.4. Agentic Safety: The New Risk Surface

The introduction of autonomous AI agents profoundly complicates the safety discussion by introducing novel, high-impact threat vectors into the T&E environment itself. Agentic systems are designed to interact autonomously with external applications, executing commands and retrieving information through function calls and structured APIs [25]. This autonomy, while powerful, dramatically magnifies the consequences of compromise.

The paramount risk is **Indirect Prompt Injection (IPI)**. In this attack vector, the malicious instruction is not provided directly by the user but is hidden within untrusted external data—such as a malicious image, a link on a webpage the agent summarizes, or an infected email message. The LLM agent processes this untrusted content, which contains instructions designed to hijack the model's output, compelling it to ignore its original directions and execute the injected malicious instruction. The consequence of a successful IPI attack on a T&E agent is severe. It can compel the agent to disclose sensitive information about the system infrastructure, provide unauthorized access to connected systems, or, most critically, automatically exfiltrate sensitive data without any user interaction [27]. A critical real-world example is the EchoLeak exploit (CVE-2025-32711) against Microsoft Copilot in mid-2025, where infected email messages containing engineered prompts could trigger the system to automatically exfiltrate data [28].

This structural flaw mandates that organizations using autonomous agents must incorporate a regulatory compulsory phase of red teaming the testing agent itself. Adopting an autonomous AI agent shifts the organization's risk exposure from addressing simple system vulnerabilities to managing complex, high-impact agentic safety risks [25]. If a network-testing agent is compromised via IPI, the security tool itself instantly transforms into a high-level threat vector—a sophisticated back-door capable of mass data exfiltration. Consequently, agent evaluation must include specific benchmarks for prompt success rates and ethical/responsible AI metrics. Defenses must employ a combination of probabilistic defense techniques (e.g., hardened system prompts, input isolation via

Spotlighting) and, where possible, deterministic blocking methods to ensure guaranteed security despite the underlying probabilistic nature of the LLM. Human in the loop- HITL research ties its origin back to Wizard-of-Oz experiments but has re-gained its popularity in the GenAI era often described as "AI-in-the-loop" and "collaborative intelligence." Frameworks like Thinker's state-machine-augmented generation have shown 82% success on realistic customer scenarios, outperforming LLM-only baselines (Thinker Framework, 2025) [29].

Enterprise Web Browser Security- The concept of securing web access within enterprise environment has evolved significantly over the past two decades, shaped by changing user behavior, threat sophistication, and infrastructure decentralization. Initially, enterprise browser security was an afterthought dependent on local antivirus software, perimeter firewalls, and user education to deter phishing and malware. However, as web-based applications became central to business operations, and remote work gained traction, the limitations of these legacy approaches became evident [30].

AI-Augmented Cybercrime-as-a-Service (CaaS)- The rise of AI is democratizing access to sophisticated cyber capabilities. No longer confined to elite nation-state actors or well-funded APT groups, the threat landscape now includes "low-tech" criminals armed with AI-powered toolkits. Simulation-based evaluation using Cloud- performance metrics derived from projections based on published AWS Lambda benchmarks rather than operational deployment. This limitation means actual performance may vary depending on factors such as network conditions, data center proximity, and real-world load patterns. [31]

Zero-Trust Architecture (ZTA) mitigates these restrictions by embracing a "never trust, always verify" principle. In Zero Trust Architecture (ZTA), trust is not predicated on network location; rather, each request—irrespective of its source—is rigorously validated, permitted, and encrypted. The fundamental concept is to regard all communications as potentially adversarial and to implement stringent security protocols at every boundary, both external and internal. [32]

4. Results

4.1 Speed and Efficiency Comparison (Velocity)

The comparison of T&E methodologies reveals a clear disparity in Velocity and scalability. Traditional methods are fundamentally constrained by human resources and time Manual Penetration testing→ offers depth but are slow and non-scalable cyber strong systems. Automated → are balanced but limited by signatures. AI/LLM T&E→ fastest, most scalable, highest zero-day performance.

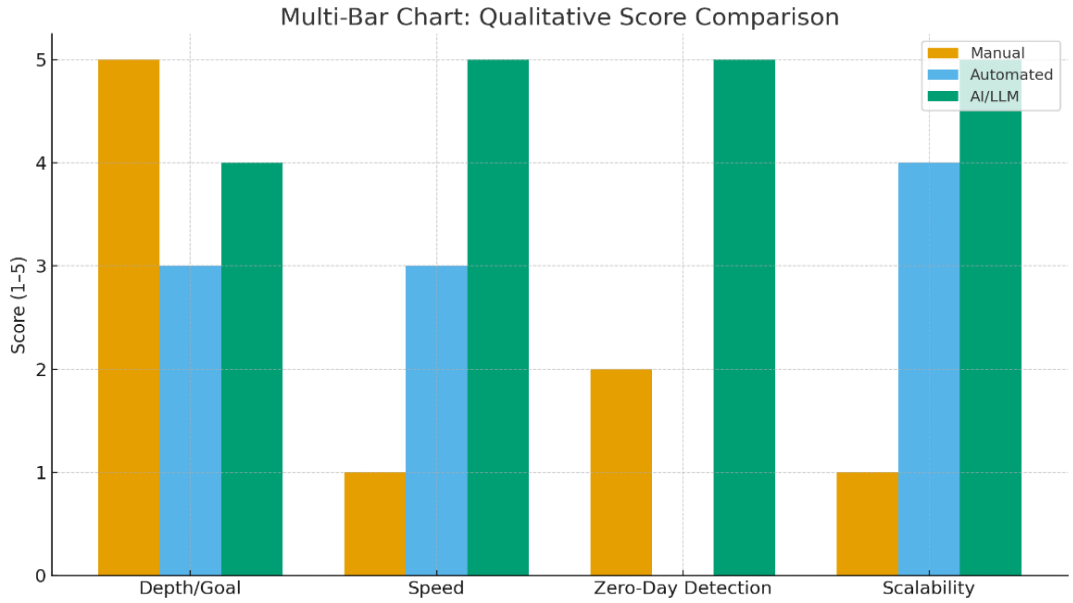

Figure 1- This chart compares all three approaches (Manual, Automated and AI/LLM) side-by-side for each metric after converting qualitative experimentation factors to numeric scores.

4.2. Safety and Accuracy Comparison (Fidelity)

Fidelity encapsulates the trustworthiness, accuracy, and depth of the security findings. In this domain, human-validated and deterministic methods maintain a critical advantage.

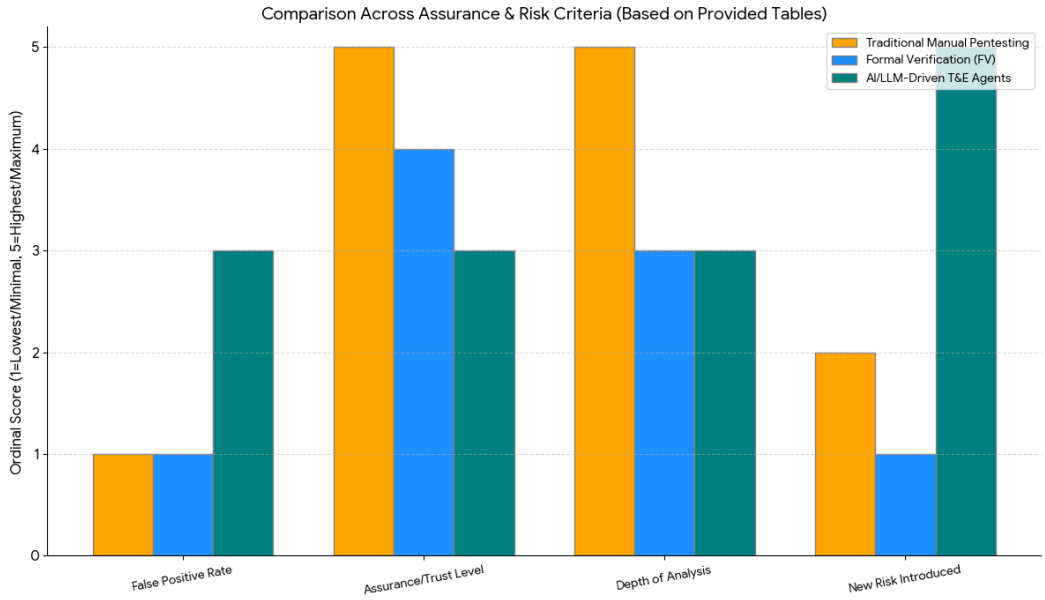

Figure 2 – Illustrates -Formal Verification (FV) dominates, followed by traditional MPT and lastly by AI-LLM Agentic security bots but they still offer high speed.

5. Conclusion

In terms of velocity and execution time taken, AI/LLM systems are definitively faster than both manual and traditional automated security methods. They provide superior threat detection by utilizing real-time machine learning and continuous data processing. AI's ability to handle millions of data points and automate repetitive tasks significantly reduces the time-to-market and accelerates the overall security delivery lifecycle. AI's ability to reduce false positives, which increases the throughput of actionable security intelligence, transforming the security team's focus from execution to strategic remediation. In terms of safer execution-Traditional methods, specifically Manual Penetration Testing and Formal Verification, are inherently safer than AI/LLM-driven T&E. Formal Verification provides deterministic assurance, while MPT delivers findings that have been manually validated and exploited, ensuring virtually zero false positives and uncovering complex logic flaws that defy algorithmic detection.

While AI improves accuracy over traditional scanners (achieving a low/moderate FPR), its probabilistic foundation, susceptibility to adversarial manipulation, agent misunderstanding, agentic manipulation and inability to move beyond "shallow" analysis limit its capacity for complex threat identification. Crucially, AI introduces significant novel risks, specifically agentic safety threats like Indirect Prompt Injection, which can compromise the security tool itself, leading to devastating outcomes such as autonomous data exfiltration. Therefore, for systems requiring the highest levels of sensitivity or trustworthiness, human oversight and deterministic verification methods and traditional methods remain indispensable and AI can safeguard less critical role applications for complete breadth and operation depth.

Refrences

1. T. Grance, T. Nolan, K. Burke, R. Dudley, G. White, and T. Good, "Special Publication 800-84 Sponsored by the Department of Homeland Security Guide to Test, Training, and Exercise Programs for IT Plans and Capabilities Recommendations of the National Institute of Standards and Technology," Sep. 2006. Available: https://nvlpubs.nist.gov/nistpubs/Legacy/SP/nistspecialpublication800-84.pdf

2. T. Zaid and S. Garai, "Emerging Trends in

Cybersecurity: A Holistic View on Current Threats, Assessing Solutions, and Pioneering New Frontiers," *Blockchain in Healthcare Today*, vol. 7, no. 1, Apr. 2024, doi: https://doi.org/10.30953/bhty.v7.302.

3. Express Computer, "AI in 5 years: Preparing for intelligent, automated cyber attacks," *Express Computer*, Dec. 02, 2025. https://www.expresscomputer.in/guest-blogs/ai-in-5-years-preparing-for-intelligent-automated-cyber-attacks/130401/ (accessed Dec. 05, 2025).

4. Vuppala, N. S. M., Gupta, D., & Yadav, S. (2025). Securing healthcare transactions in AI-augmented systems: A comprehensive framework for enhanced cybersecurity in health insurance operations. Emerging Frontiers Library for The American Journal of Applied Sciences, 7(10), 44–51.

5. "The Shortcomings of Traditional Penetration Tests -and How Autonomous Pentesting Addresses Them with Research and Analysis by SPONSORED BY." Accessed: Dec. 05, 2025. [Online]. Available: https://horizon3.ai/wp-content/uploads/2022/07/IDC_Report___The_Shortcomings_of_Traditional_Penetration_Tests.pdf

6. "AI Testing vs Traditional Testing: Pros, Cons, and the Right Fit for Your Business," *Frugaltesting.com*, 2025. https://www.frugaltesting.com/blog/ai-testing-vs-traditional-testing-pros-cons-and-the-right-fit-for-your-business (accessed Dec. 05, 2025).

7. NIST, "Guide for Conducting Risk Assessments," NIST, Sep. 2012. Available: https://nvlpubs.nist.gov/nistpubs/legacy/sp/nistspecialpublication800-30r1.pdf

8. Data Driven Trade-Off Analysis for Cybersecurity - ODU Digital Commons, accessed December 4, 2025, https://digitalcommons.odu.edu/cgi/viewcontent.cgi?article=1241&context=emse_fac_pubs

9. S. Bhattacharjee, "Static vs. dynamic code analysis: A comprehensive guide," *vFunction*, Sep. 04, 2024. https://vfunction.com/blog/static-vs-dynamic-code-analysis/

10. Bitdefender Enterprise, "The Differences Between Static and Dynamic Malware Analysis," *Bitdefender Blog*, Aug. 29, 2023. https://www.bitdefender.com/en-us/blog/businessinsights/the-differences-between-static-malware-analysis-and-dynamic-malware-analysis

11. Gunda, S. K. (2024, September). Analyzing Machine Learning Techniques for Software Defect Prediction: A Comprehensive Performance Comparison. In 2024 Asian Conference on Intelligent Technologies (ACOIT) (pp.1-5). IEEE.

12. SecureFlag, "Why Static Analysis Alone Isn't Enough," *SecureFlag*, Aug. 19, 2025. https://blog.secureflag.com/2025/08/19/why-static-analysis-isnt-enough/

13. A. Pahuja, "Automated vs Manual Penetration Testing: Which One You Need," *www.getastra.com*, Mar. 15, 2022. https://www.getastra.com/blog/security-audit/automated-vs-manual-penetration-testing/

14. M. Khalil, "Manual vs Automated Penetration Testing: The 2025 Guide," *DeepStrike*, Aug. 09, 2025. https://deepstrike.io/blog/manual-vs-automated-penetration-testing

15. P. S. Shinde and S. B. Ardhapurkar, "Cyber security analysis using vulnerability assessment and penetration testing," *2016 World Conference on Futuristic Trends in Research and Innovation for Social Welfare (Startup Conclave)*, Coimbatore, India, 2016, pp. 1-5, doi: 10.1109/STARTUP.2016.7583912.

16. Vinugayathri Chinnasamy, "Manual vs Automated Penetration Testing: When & Why to Use," *Indusface*, Jul. 31, 2025. https://www.indusface.com/blog/manual-vs-automated-pen-testing/

17. S. K. Gunda, "Automatic Software Vulnerabilty Detection Using Code Metrics and Feature Extraction," 2025 2nd International Conference On Multidisciplinary Research and Innovations in Engineering (MRIE), Gurugram, India, 2025, pp. 115-120, https://doi.org/10.1109/MRIE66930.2025.11156601

18. P. R. Rajgopal, "AI Threat Countermeasures: Defending Against LLM-Powered Social Engineering," *International journal of IoT*, vol. 5, no. 02, pp. 23–43, Aug. 2025, doi: https://doi.org/10.55640/ijiot-05-02-03

19. Formal verification: how a 400-year-old mathematical idea could transform cybersecurity, accessed December 4, 2025, https://www.thalesgroup.com/en/news-centre/insights/formal-verification-how-400-year-old-mathematical-idea-could-transform

20. New Study | AI Outperforms Humans in Detecting Zero-Day Threats - Cyber Security, accessed December 4, 2025, https://www.cybersecurityinstitute.in/blog/new-study-ai-outperforms-humans-in-detecting-zero-day-threats

21. Gunda, Sai Krishna & Yalamati, Srinivasu & Gudi, Srikanth Reddy & Manga, Indrasena & Aleti, Akhilesh Kumar. (2025). SCALABLE AND ADAPTIVE MACHINE LEARNING MODELS FOR EARLY SOFTWARE FAULT PREDICTION IN AGILE DEVELOPMENT: ENHANCING SOFTWARE RELIABILITY AND SPRINT PLANNING EFFICIENCY. 10.12732/ijam.v38i2s.74.

22. [2509.07933] Breaking Android with AI: A Deep Dive into LLM-Powered Exploitation - arXiv, accessed December 4, 2025, https://arxiv.org/abs/2509.07933

23. [2410.17141] Towards Automated Penetration Testing: Introducing LLM Benchmark, Analysis, and Improvements - arXiv, accessed December 4, 2025, https://arxiv.org/abs/2410.17141

24. LLMs Evaluation: Benchmarks, Challenges, and Future Trends - Prem AI, accessed December 4, 2025, https://www.premai.io/blog/llms-evaluation-benchmarks-challenges-and-future-trends

25. AI Privacy Risks & Mitigations – Large Language Models (LLMs) - European Data Protection Board, accessed December 4, 2025, https://www.edpb.europa.eu/system/files/2025-04/ai-privacy-risks-and-mitigations-in-llms.pdf

26. The Ethics of Interactions: Mitigating Security Threats in LLMs - arXiv, accessed December 4, 2025, https://arxiv.org/html/2401.12273v2

27. LLM01:2025 Prompt Injection - OWASP Gen AI Security Project, accessed December 4, 2025, https://genai.owasp.org/llmrisk/llm01-prompt-injection/

28. Agentic AI Security: Threats, Defenses, Evaluation, and Open Challenges - arXiv, accessed December 4, 2025, https://arxiv.org/html/2510.23883v1

29. Vuppala, N. S. M. (2025). Human-in-the-loop and generative AI dilemma: A hybrid strategy for effective customer service in enterprise CRM. International Journal of Business and Technology Management. https://mysitasi.mohe.gov.my/journal-website/journal/ijbtm/

30. Prassanna Rao Rajgopal . Cybersecurity Platformization: Transforming Enterprise Security in an AI-Driven, Threat-Evolving Digital Landscape. International Journal of Computer Applications. 186, 80 (Apr 2025), 19-28. https://doi.org/10.5120/ijca2025925611

31. Pagidoju, R. T., & Agarwal, S. (2025, October). Cloud-Native Generative AI for Automated Planogram Synthesis: A Diffusion Model Approach for Multi-Store Retail Optimization. In *International Conference on Software Engineering and Data Engineering* (pp. 152-165). Cham: Springer Nature Switzerland.

32. Sagar Kesarpu. (2025). Zero-Trust Architecture in Java Microservices. *International Journal of Networks and Security*, *5*(01), 202–214. Retrieved from https://inlibrary.uz/index.php/ijns/article/view/110609