Article 2025 | Author Nakul Marwah
Nakul Marwah | Nakul.marwah@gmail.com

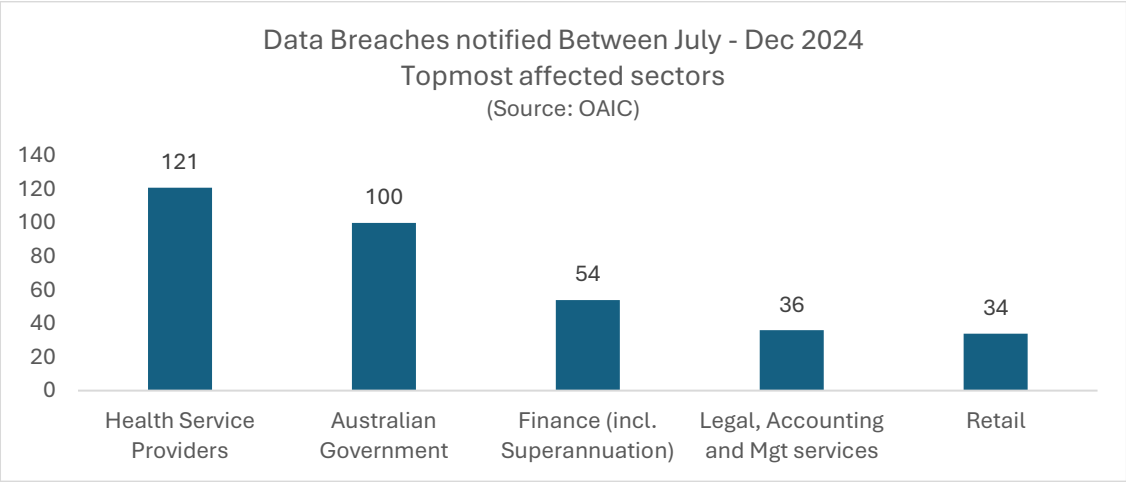
The Double-Edged Sword: AI's Transformative Impact on Vulnerability Management

How artificial intelligence is simultaneously revolutionising and complicating the cybersecurity landscape

Introduction

The cybersecurity landscape is evolving at an unprecedented pace, with artificial intelligence emerging as both a powerful ally and a formidable adversary in vulnerability management. As organisations grapple with an ever-expanding attack surface and increasingly sophisticated threats, AI technologies present a paradox: they offer revolutionary capabilities for identifying and mitigating vulnerabilities while simultaneously introducing new classes of risks that traditional security frameworks struggle to address.

The rise of vulnerabilities and consequential impact is particularly acute in Australia. There was a 15% increase in reported data breaches in the latter half of 2024 compared with the first half [(1)]. The most affected sectors (in decreasing order) were: health service providers, the Australian government, Finance, Professional services, and Retail [(1)]. According to ASD's Annual Cyber Threat Report 2023 - 24, 34% of respondents reported that their financial or personal information was exposed in a data breach within 12 months [(2)]. Given this heightened risk environment, it is concerning that 36% of Australian cyber security leaders and practitioners, more than twice that of global peers, strongly agree that their organisations are not adequately prepared to defend against AI-powered threats and attacks [(3)].



For technology governance professionals, understanding this duality is crucial. The integration of AI into vulnerability management processes represents one of the most significant paradigm shifts in cybersecurity since the advent of automated threat detection. Yet, as we embrace these powerful tools, we must also confront the uncomfortable reality that AI systems

themselves have become critical assets requiring protection—and potential vectors for exploitation.

The AI Advantage: Transforming Traditional Vulnerability Management

Automated Discovery and Prioritisation

Traditional vulnerability management has long been hampered by the sheer volume of potential security issues that organisations must address. The average enterprise discovers thousands of vulnerabilities annually, highlighting the complexity of accurately identifying high-risk vulnerabilities and creating a prioritisation nightmare that often leaves critical exposures unaddressed.

AI-powered vulnerability management platforms are revolutionising this process through intelligent automation. Machine learning algorithms can analyse vast datasets of vulnerability information, correlating threat intelligence feeds with organisational context to provide risk-based prioritisation that goes far beyond simple CVSS scores. These systems consider factors such as asset criticality, network topology, active threat campaigns, and historical exploit patterns to create dynamic risk assessments.

Further, AI excels at pattern recognition, enabling the identification of vulnerability clusters and attack paths that human analysts or classic tools might miss. By analysing code repositories, network configurations, and system dependencies, AI tools can predict potential cascade effects and identify vulnerabilities that pose the greatest systemic risk to an organisation.

Accelerated Remediation and Response

The speed advantage of AI in vulnerability management cannot be overstated. While classic mechanisms may take days or weeks to fully assess and respond to new vulnerabilities, AI systems can process and act on threat intelligence in near-real-time. Automated patch management systems powered by machine learning can evaluate patch compatibility, schedule deployments during optimal windows, and even perform rollback operations if issues are detected.

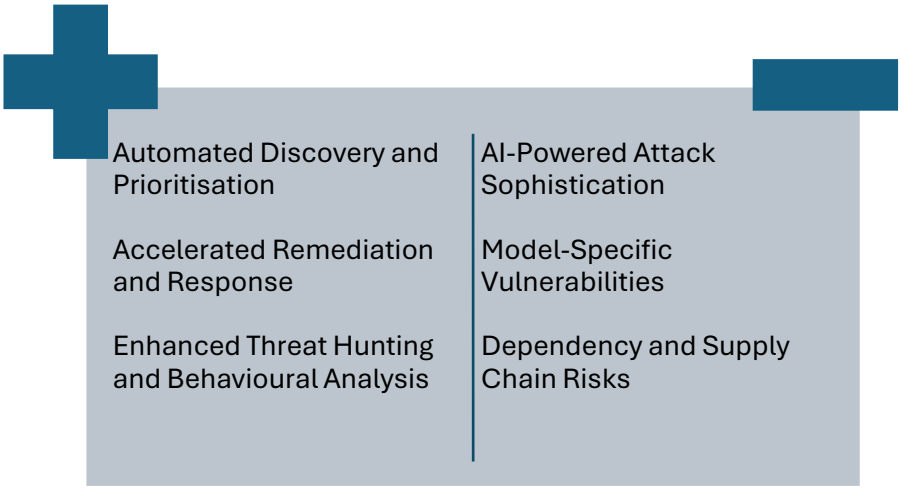
AI also enables more sophisticated compensating controls. When immediate patching isn't feasible, AI systems can automatically implement network segmentation rules, adjust firewall configurations, or deploy virtual patches through web application firewalls—all while learning from each deployment to improve future responses.

Enhanced Threat Hunting and Behavioural Analysis

Perhaps most significantly, AI is transforming vulnerability management from a reactive discipline into a predictive one. Advanced behavioural analytics can identify subtle indicators of compromise that suggest vulnerability exploitation, even when traditional signatures fail to

detect malicious activity. These systems establish baselines of normal behaviour and flag anomalies that might indicate successful attacks. In the Australian context, this has become increasingly critical, as CyberCX's threat report highlights that, in 2024, espionage-related incidents took more than two weeks longer to detect than in 2023. [4]

This capability extends to zero-day detection, where AI models trained on vast datasets of attack patterns can identify novel exploitation techniques and potentially vulnerable code patterns before they're formally documented as CVEs.



The Dark Side: AI-Enabled Risks and New Attack Vectors

AI-Powered Attack Sophistication

While AI strengthens defensive capabilities, it simultaneously empowers adversaries with unprecedented offensive capabilities. AI-enhanced vulnerability scanners can now perform reconnaissance at scale, automatically identifying and exploiting vulnerabilities across vast networks with minimal human intervention. These tools can adapt their tactics in response to defensive measures, making them particularly challenging to counter with traditional security measures.

More concerning is the emergence of AI systems capable of generating novel exploits. By training on large datasets of vulnerability disclosures and exploit code, these systems can potentially identify new attack vectors or create exploits for recently discovered vulnerabilities. Traditional tools are not capable enough to defend against such AI-powered attacks. Only 30% of Australian professionals have shown confidence in the ability of traditional cybersecurity solutions to detect and block AI-powered threats and attacks [3].

Model-Specific Vulnerabilities

AI systems introduce entirely new categories of vulnerabilities that traditional security frameworks don't adequately address. Adversarial attacks against machine learning models can cause them to misclassify threats, ignore genuine vulnerabilities, or incorrectly prioritise

risks. Data poisoning attacks can corrupt training datasets, compromising decision-making across entire vulnerability management programs.

Model theft and extraction attacks represent another significant concern. If adversaries can reverse-engineer the AI models used in vulnerability management systems, they can identify blind spots and develop evasion techniques specifically tailored to those systems.

Dependency and Supply Chain Risks

The integration of AI into vulnerability management creates new dependencies and potential single points of failure. Organisations increasingly rely on third-party AI services, cloud-based machine learning platforms, and pre-trained models, each of which represents a potential attack vector. A compromise of these upstream dependencies could have cascading effects across multiple organisations' security postures.

This concern is particularly relevant for Australian organisations, where supply chain vulnerabilities have been identified as a critical risk factor. In July 2023 – June 2024, the Australian Signal Directorate (ASD) responded to 107 cyber supply chain incidents, comprising of 9% of all cybersecurity incidents responded to by ASD [(2)].

Additionally, the complexity of AI systems makes them difficult to audit and validate. Unlike traditional software vulnerabilities that can be identified through code review, AI model vulnerabilities may emerge only under specific conditions or when exposed to particular input data.

Strategic Mitigation Approaches

Implementing AI Governance Frameworks

Organisations must develop comprehensive AI governance frameworks that address both the benefits and risks of AI integration in vulnerability management. This includes establishing clear policies for AI model validation, regular auditing of AI-driven decisions, and maintaining human oversight of critical security functions.

Effective governance also requires cross-functional collaboration between security teams, data scientists, and risk management professionals. Vulnerability management can no longer be viewed solely as a technical function; it requires ongoing strategic oversight that accounts for the broader implications of AI integration.

Diversification and Resilience Strategies

To mitigate the risks associated with AI dependency, organisations should adopt diversified approaches that combine multiple AI models, maintain human expertise, and preserve manual override capabilities. This defence-in-depth strategy ensures that the failure or compromise of any single AI system doesn't create critical gaps in vulnerability management capabilities.

Regular red team exercises specifically focused on AI system vulnerabilities can help identify potential weaknesses before they're exploited by adversaries. These exercises should include attempts to poison training data, evade model detection, and exploit AI-specific vulnerabilities.

Continuous Monitoring and Adaptation

The dynamic nature of AI requires continuous monitoring and adaptation of vulnerability management processes. Organisations must establish metrics and monitoring systems that can detect when AI models are performing sub optimally or potentially being manipulated. This includes tracking model accuracy over time, monitoring for unusual patterns in vulnerability detection, and maintaining baseline performance metrics.

The Path Forward: Balanced Integration

As we navigate this new landscape, the key to success lies not in choosing between AI and traditional approaches, but in thoughtfully integrating both while remaining cognizant of the associated risks. Organisations that can effectively harness AI's power while implementing robust safeguards will gain significant competitive advantages in vulnerability management.

The future of cybersecurity will be defined by our ability to manage this balance. As AI continues to evolve, so too must our approaches to governance, risk management, and security architecture. The organisations that invest now in understanding and addressing AI-related vulnerabilities while leveraging AI's defensive capabilities will be best positioned to thrive in an increasingly complex threat landscape.

The double-edged nature of AI in vulnerability management is not a problem to be solved, but a reality to be managed. By embracing this complexity and building adaptive, resilient security programs, we can harness the transformative power of AI while mitigating its potential pitfalls.

---

(1) Notifiable Data Breaches Report: July to December 2024 | OAIC

(2) Annual Cyber Threat Report 2023-2024 | Cyber.gov.au

(3) State of Cybersecurity AI (2024) - Australia Edition | Resources | Darktrace

(4) CyberCX 2025 Threat Report Highlights Emerging Risks