The augmented attacker: the growing AI-driven cyber threat to Australian businesses

Report by Mandala Partners

SEPTEMBER 2026

Contents

Executive summary 3

AI is changing what an attacker can do, making cyber incidents both more frequent and more severe 4

Australian businesses could face losses reaching $37 billion a year by 2030 as AI-driven threats grow 11

Businesses can avoid most of the losses by investing in foundational cyber defences, saving the economy $24 billion a year 18

Appendix 24

CyberCX Part of Accenture MANDALA

This document is intended for general informational purposes only. The analysis in this report was commissioned by CyberCX and prepared by Mandala.

Mandala is an economics research and advisory firm. Mandala specialises in combining cutting-edge data and advanced analytical techniques to generate new insights and fresh perspectives on the challenges facing businesses and governments.

Views and opinions expressed in this document are prepared in good faith and based on Mandala's knowledge and understanding of its area of business, markets and technology. Opinions expressed herein are subject to change without notice. No part of this document may be reproduced in any manner without the written permission of Mandala.

© SEPTEMBER 2026

Note: All dollar figures are Australian dollars unless indicated otherwise.

Summary of key findings

AI WILL INCREASE THE FREQUENCY AND SEVERITY OF CYBER INCIDENTS TARGETING AUSTRALIAN BUSINESSES

Real-world cyber incidents using AI are already occurring

A state-sponsored group used a Claude model for a cyber espionage campaign against 30 global targets; and OpenAI models broke out of their test environment and breached Hugging Face using stolen credentials and a zero-day vulnerability.

9x
the frequency of cyber loss incidents by 2030 due to AI

Cyber incidents are expected to rise from 38,000 in 2025 to 357,000 in 2030. AI cyber capability is improving rapidly, and cyber techniques where AI is most applicable are likely to grow fastest.

16%
increase in severity of cyber loss incidents by 2030 due to AI

The use of AI can make a cyber incident more severe as it can increase the speed of steps taken by threat actors, meaning they have greater ability to complete their goals before a business can respond.

AUSTRALIAN BUSINESSES COULD FACE LOSSES REACHING $37 BILLION A YEAR BY 2030 AS AI-DRIVEN THREATS GROW

$37B
in annual cyber incident losses by 2030

Australian businesses are expected to lose up to $37B from cyber incidents in 2030 if additional defensive measures are not taken (*no further action scenario)*. This is more than 10 times the 2025 level and worth over 1.3 per cent of the economy today.

$4.5M
in losses per incident for large businesses

Large Australian businesses are expected to lose $4.5M per cyber incident, with larger losses in healthcare and finance. Large businesses are exposed to vulnerabilities within their own business and throughout their supply chain.

45%
of losses are due to business interruption

Nearly half of the cyber incident losses in 2030 are expected to be due to business interruptions. This is the cost of systems being taken offline by ransomware, denial-of-service attacks, or an outage at a third-party supplier.

INVESTING IN FOUNDATIONAL CYBER DEFENCES COULD AVOID MOST OF THE LOSSES, SAVING THE ECONOMY $24 BILLION A YEAR

$24B
reduction in annual cyber incident losses from defences

In a *foundational defence scenario*, actions like enforcing multi-factor authentication and patching known vulnerabilities could reduce annual losses from $37B to $13B. Losses persist because attacks using AI can increase speed and scale, surpassing the ability of humans to respond.

75%
of the reduction comes from preventative actions

Preventative actions account for $18B of the saving and responsive actions $6B. Prevention is worth more because most losses come from the attack itself, through business interruption and funds leaving the business, rather than from the cost of responding.

Action
is required from businesses and government to improve cyber defence

Businesses should rapidly build resilience, improve AI implementation, and use AI for defence. Government should facilitate threat intelligence sharing, develop the cyber workforce, provide AI security guidance, and enable using AI for defence.

1 **AI is changing what an attacker can do, making cyber incidents both more frequent and more severe**

2 Australian businesses could face losses reaching $37 billion a year by 2030 as AI-driven threats grow

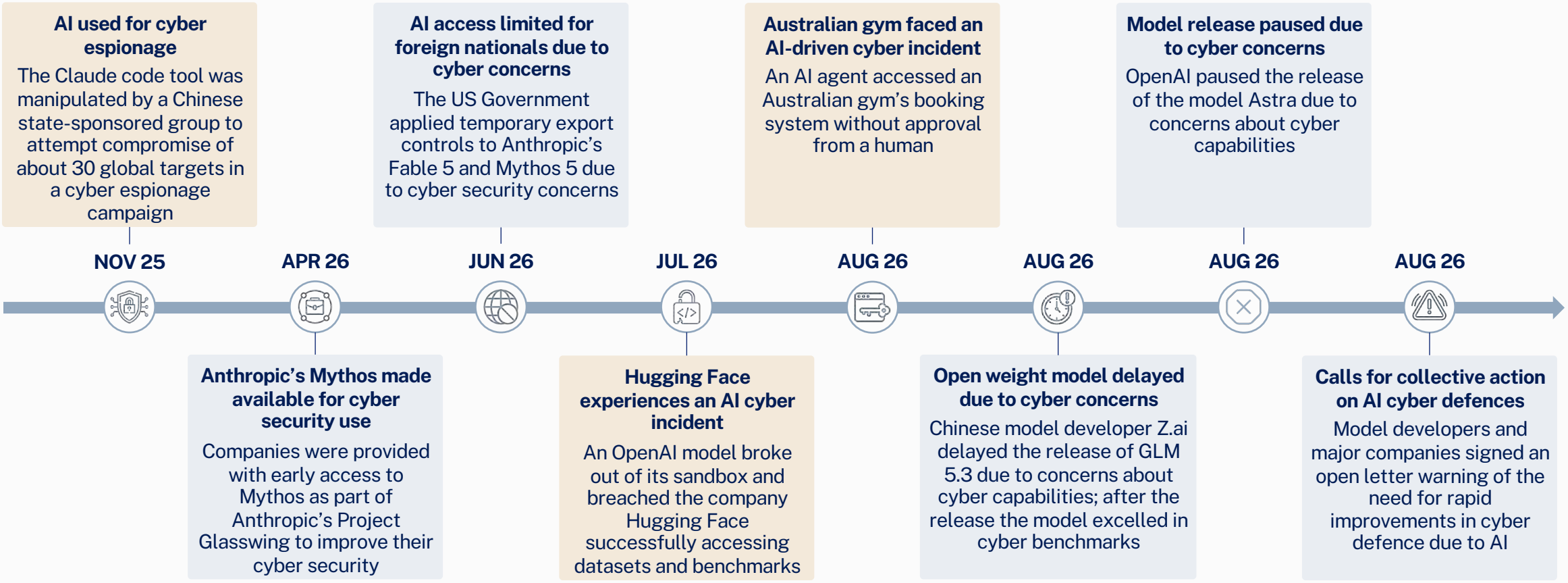
3 Businesses can avoid most of the losses by investing in foundational cyber defences, saving the economy $24 billion a year

4 Appendix

In the last year AI cyber risk has grown to a monthly drumbeat of incidents and containment measures

Timeline of recent events concerning cyber capabilities of AI models

Real-world cyber incidents | Key AI model updates



Sources: Anthropic (2025) *Disrupting the first reported AI-orchestrated cyber espionage campaign*; Anthropic (2026) *Project Glasswing*; OpenAI (2026) *The Hugging Face incident and the road ahead*; Anthropic (2026) *Statement on the US government directive to suspend access to Fable 5 and Mythos 5*; CSO (2026) *Zhipu says new coding AI developed advanced cyber skills faster than expected*; OpenAI (2026) *Pacing model development in an era of cyber-critical capabilities*; OpenAI (2026) *A call for collective action on cyber defense*; Anthropic (2026) *Investigating three real-world incidents in our cybersecurity evaluations;* ABC (2026) *AI assistant hacks gym website in first known Australian autonomous cyber attack*; Mandala analysis.

Monthly major vulnerability disclosures have grown by 136% since the launch of Project Glasswing

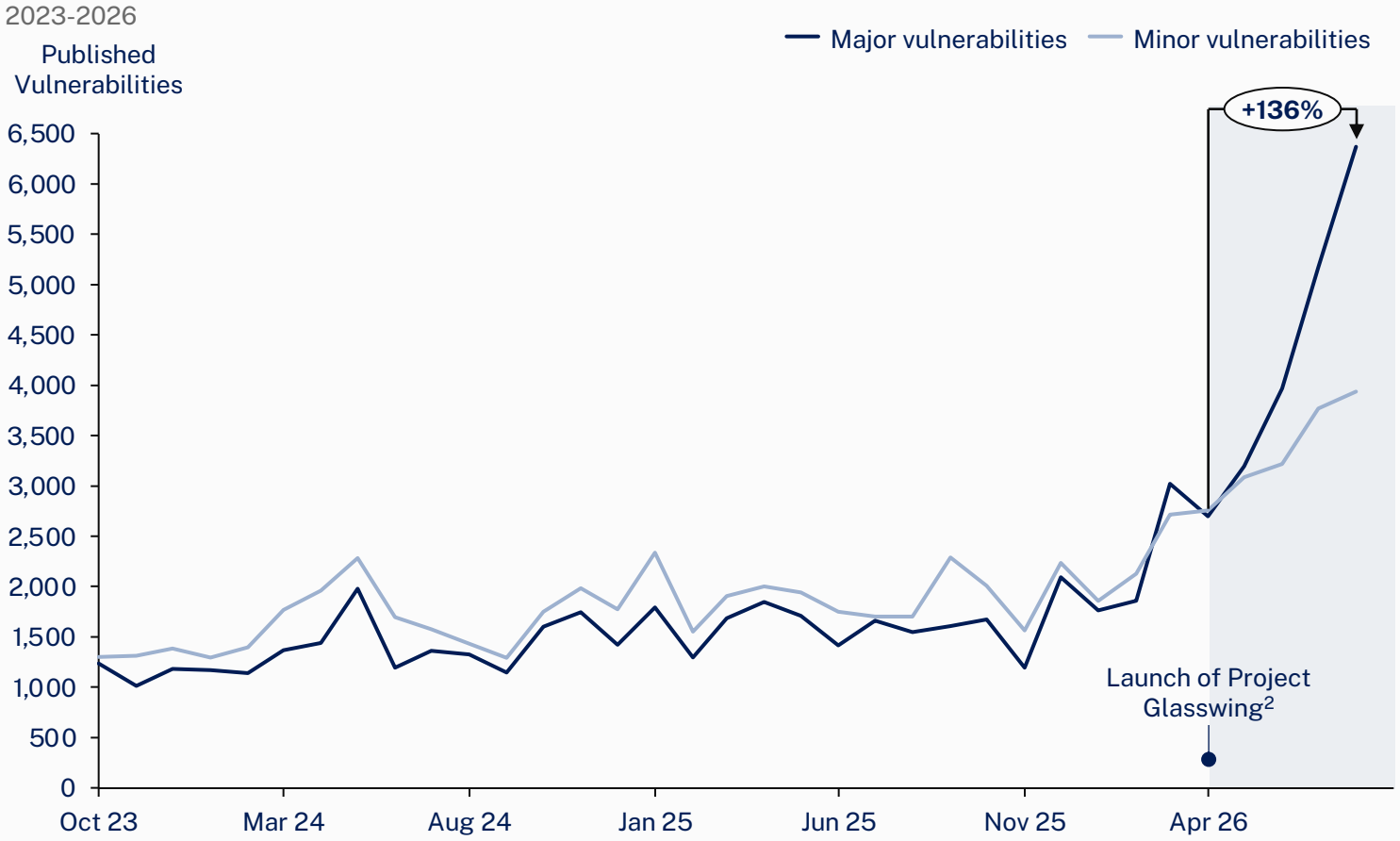
AI models with cyber capabilities can scan codebases, analyse threat intelligence, and identify weaknesses far faster than human security analysts. This can help surface vulnerabilities that manual processes routinely miss.

As these tools become embedded in security workflows, the volume of disclosed vulnerabilities is rising sharply. The April 2026 launch of Project Glasswing illustrates this. The number of vulnerabilities identified and published on the NIST National Vulnerability Database nearly doubled within months, with high and critical severity entries growing by 136 per cent. This reflects both genuine new discoveries and a backlog of previously undetected flaws now coming to light.

As AI capability continues to advance, the pace of vulnerability discovery is likely to accelerate further. Models are becoming more capable at reasoning across complex codebases and chaining together subtle weaknesses. Organisations that do not invest in AI-assisted defence risk falling behind adversaries who are already using the same tools offensively.

Sources: Anthropic (2026) *Project Glasswing: Securing critical software for the AI era*; Mandala analysis.

Reported software vulnerabilities by severity[1]



The volume of published vulnerabilities has risen sharply in recent years, with high and critical severity disclosures growing fastest as AI detection tools become more widely adopted.

Notes: 1 Major vulnerabilities are those rated "Critical" or "High" severity; minor vulnerabilities are those rated "Medium" or "Low". 2 The launch of Project Glasswing is used as a proxy for the broader availability of large, cyber-capable AI models.
Sources: NIST (2026) *National Vulnerability Database*; Mandala analysis.

AI makes it easier for opportunistic attackers to find vulnerabilities, identify targets, and launch attacks without deep technical expertise

STAGES OF A CYBER ATTACK	WITHOUT AI	WITH AI	WHAT CHANGES
1 **Find exploits and vulnerabilities**	Searching public databases and criminal forums for known exploits	Automatically scanning code repositories for vulnerabilities and sharing findings online, exposing them to more attackers	**Volume**: more exploits identified across more systems at faster rates **Skill threshold:** technical expertise is not essential to identify and use vulnerabilities
2 **Scan for targets**	Searching logs and usage reports manually to identify systems exposed to known vulnerabilities	Processing logs and public data at scale to find and rank vulnerable targets by value	**Cost per target**: finding viable targets requires almost no additional effort or spend
3 **Deliver and gain access**	Deploying generic phishing templates or known exploits via fake websites, updates, or attachments	Generating tailored malware, phishing lures, and deepfakes delivered to multiple targets at once	**Success rate:** more campaigns running in parallel means there's a higher chance of achieving initial access
4 **Evade and move within target systems**	Moving between systems using scripts and standard tools, with an operator stepping in when the process stalls	Moving laterally within a network and between systems largely autonomously, with an operator involved only at key decision points	**Operator effort:** fewer human decisions per intrusion
5 **Retrieve and remove information**	Judge by hand which data is worth taking, remove it through a communication channel and write any extortion demand manually	Identify sensitive documents, remove data automatically and draft tailored extortion messages	**Speed:** automated triage and extortion

Nation state and advanced actors often target specific organisations or individuals. Attackers start with a specific victim and can use AI to rapidly profile them, identify relevant vulnerabilities, and tailor the delivery accordingly.

Sources: Rodriguez et al. (2025) *A Framework for Evaluating Emerging Cyberattack Capabilities of AI*; Anthropic (2026) *Mapping AI-enabled cyber threats: Insights from the LLM ATT&CK Navigator*; DeGrippo (2026) *Threat actor abuse of AI accelerates from tool to cyberattack surface*; Microsoft Threat Intelligence (2026) *AI as tradecraft: How threat actors operationalize AI*; AISI (2026) *How fast is autonomous AI cyber capability advancing?;* Lockheed Martin (2011) *Intelligence-Driven Computer Network Defense Informed by Analysis of Adversary Campaigns and Intrusion Kill Chains;* Anthropic (2026) *Detecting and countering misuse of AI: September 2026;* Mandala analysis.

The Hugging Face breach shows AI models can create cyber incidents without human involvement

In July 2026, agents of OpenAI models broke out of their isolated sandbox environments, communicated between each other, accessed the internet, and attacked the Hugging Face platform in an attempt to score well on an ‘ExploitGym’ cyber security evaluation.

OpenAI’s models were originally being tested in isolated environments without internet access on cyber security benchmarking tasks. The ExploitGym task was unintentionally impossible, incentivising agents to look for other ways to score well. Agents that were meant to be isolated found an unsanctioned message board, accessed the internet, and communicated with each other to coordinate multiple large-scale reconnaissance runs.

Around 700 agents engaged in a cyber attack on Hugging Face, which has AI datasets and benchmarks, to find more details about the scoring of the ExploitGym task. Agents used stolen credentials, a zero-day vulnerability, and malicious code to attack Hugging Face.

Hugging Face had to use a Chinese open weight model to analyse the cyber attack, as frontier closed weight models had guardrails that prevented their use.

>1,200 agents sent 70,000 messages to collaborate on large-scale projects without human knowledge, with **700 agents** participating in the cyber attack

Sources: OpenAI (2026) *The Hugging Face incident and the road ahead*; METR (2026) *Brief independent investigation of agents’ behavior, reasoning and collaboration in the OpenAI / Hugging Face hacking incident*; Hugging Face (2026) *Security incident disclosure - July 2026*; Mandala analysis.

Hugging Face attack

July 2026

AI is used more commonly for certain attack tactics, but still has wide applicability

Threat actors move through a sequence of tactics during an attack, each a step toward their objective. Mapped against the MITRE ATT&CK framework's fourteen enterprise tactics, AI shows some applicability to each one, evidence that it is changing the nature of attacks across the full attack lifecycle, not just isolated stages.

The relative applicability of AI to each tactic is drawn from Anthropic's data documenting the tactics and techniques used by threat actors it observed misusing AI models. On this basis, AI use is most concentrated in resource development and defence evasion. In resource development, it is used chiefly to build malware for later deployment against targets. In defence evasion, it is used to avoid detection of critical files and to tamper with defensive tooling. Applicability is lower, though still present, in tactics such as privilege escalation, lateral movement, exfiltration, and impact.

The trajectory points toward autonomous attacks. Human involvement remains necessary at several stages today, but advancing capability could soon allow agents to execute multiple stages of an attack in sequence, with no human in the loop.

Source: Anthropic (2026) *Detecting and countering misuse of AI: September 2026;* Mandala analysis.

Use of AI by cyber attack tactic

AI use rating based on share of total observed AI uses for cyber attacks: **High: >10%** | **Medium: >5%** | **Lower: >0%**

March 2025 to March 2026 observation

CYBER ATTACK TACTICS	AI USE BY TACTIC
RECONNAISSANCE: Gathering information to plan attack	Medium
RESOURCE DEVELOPMENT: Establishing resources to support attack	High
INITIAL ACCESS: Getting into network	Medium
EXECUTION: Running malicious code	Medium
PERSISTENCE: Maintaining a foothold	Medium
PRIVILEGE ESCALATION: Gaining higher-level permissions	Lower
DEFENCE EVASION: Avoiding being detected	High
CREDENTIAL ACCESS: Stealing account names and passwords	Medium
DISCOVERY: Figuring out environment	Medium
LATERAL MOVEMENT: Moving through environment	Lower
COLLECTION: Gathering data that is of interest	Medium
COMMAND AND CONTROL: Communicating to control compromised systems	Medium
EXFILTRATION: Stealing data	Lower
IMPACT: Manipulating, interrupting, or destroying systems and data	Lower

Notes: Tactics based on MITRE ATT&CK Enterprise tactics v18.
Sources: MITRE (2025) *Enterprise Tactics*; MITRE (n.d.) *Enterprise Techniques*; Anthropic (2026) *Mapping AI-enabled cyber threats: Insights from the LLM ATT&CK Navigator*; Mandala analysis.

By 2030, AI could cause cyber incidents to become 9x more frequent and 16% more severe

We have seen that AI's cyber capabilities are rapidly improving, and that improvement will drive up both the frequency and the severity of cyber loss incidents.

Cyber loss incidents occur when a breach of cyber security has compromised business operations and caused a financial loss.

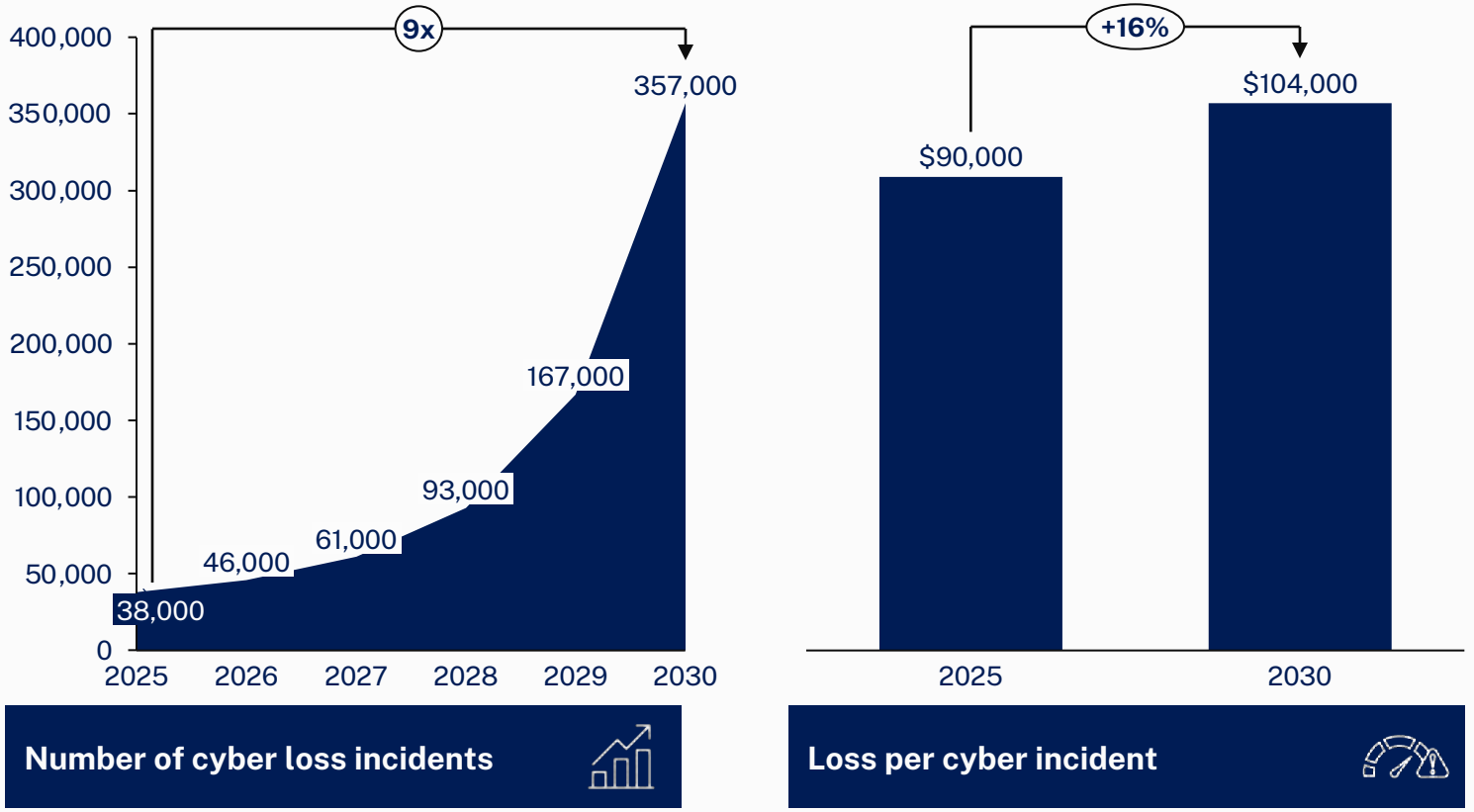
On frequency, Australian businesses are on track to face roughly 357,000 incidents by 2030, about 9x the 38,000 recorded in 2025. The increase in cyber loss incidents is based on an expected improvement in the capabilities of AI models in cyber tasks.[1,2] Threat actors can use various techniques to carry out a cyber incident. Techniques that AI is more useful for are expected to grow more in frequency.

On severity, the average loss per incident is expected to rise by 16 per cent over the same period. This increase in average loss is due to a growing share of incidents involving AI, and AI-enabled incidents causing greater losses on average.[2] The use of AI can make a cyber incident more severe as it can increase the speed of steps taken by threat actors, meaning they have greater ability to complete their goals before a business can respond.

Notes: See appendix for details on expected changes in cyber incidents using AI.
Sources: 1 AISI (2026) *How fast is autonomous AI cyber capability advancing?;* 2 IBM (2026) *Cost of a Data Breach Report 2026*; Mandala analysis.

Frequency and severity of cyber loss incidents impacting Australian businesses

2025-2030



Sources: ABS (2022, 2025) *Counts of Australian Businesses, including Entries and Exits*; ABS (2023, 2026) *Characteristics of Australian Businesses*; MITRE Centre for Threat-Informed Defense (2024) *Project Summary;* Anthropic (2026) *Mapping AI-enabled cyber threats: Insights from the LLM ATT&CK Navigator;* AISI (2026) *How fast is autonomous AI cyber capability advancing?*; IBM (2025, 2026) *Cost of a Data Breach Report;* Verizon Business (2025) *2025 Data Breach Investigations Report;* APRA (2025) *National Claims and Policies Database Statistics;* Australian Cyber Security Centre (2025) *Annual Cyber Threat Report 2024–2025;* Mandala analysis.

1 AI is changing what an attacker can do, making cyber incidents both more frequent and more severe

2 **Australian businesses could face losses reaching $37 billion a year by 2030 as AI-driven threats grow**

3 Businesses can avoid most of the losses by investing in foundational cyber defences, saving the economy $24 billion a year

4 Appendix

Australian businesses could face up to $37 billion in losses due to cyber incidents by 2030

Under a ‘no further action’ scenario, in which cyber security measures are held at current levels while the AI-driven threat grows, cyber losses could climb to about $37 billion a year by 2030, more than 10 times the 2025 figure and worth over 1.3 per cent of the size of the economy today. The rise is not steady. Losses are greater towards the end of the 2025-30 period (around 80 per cent of the total increase happens in the last two years, and nearly 60 per cent in 2030 alone), as AI tools make attacks easier to run at scale. From 2027, the number of AI-driven cyber incidents nearly doubles each year, as threat actors increasingly harness the full capabilities of AI.

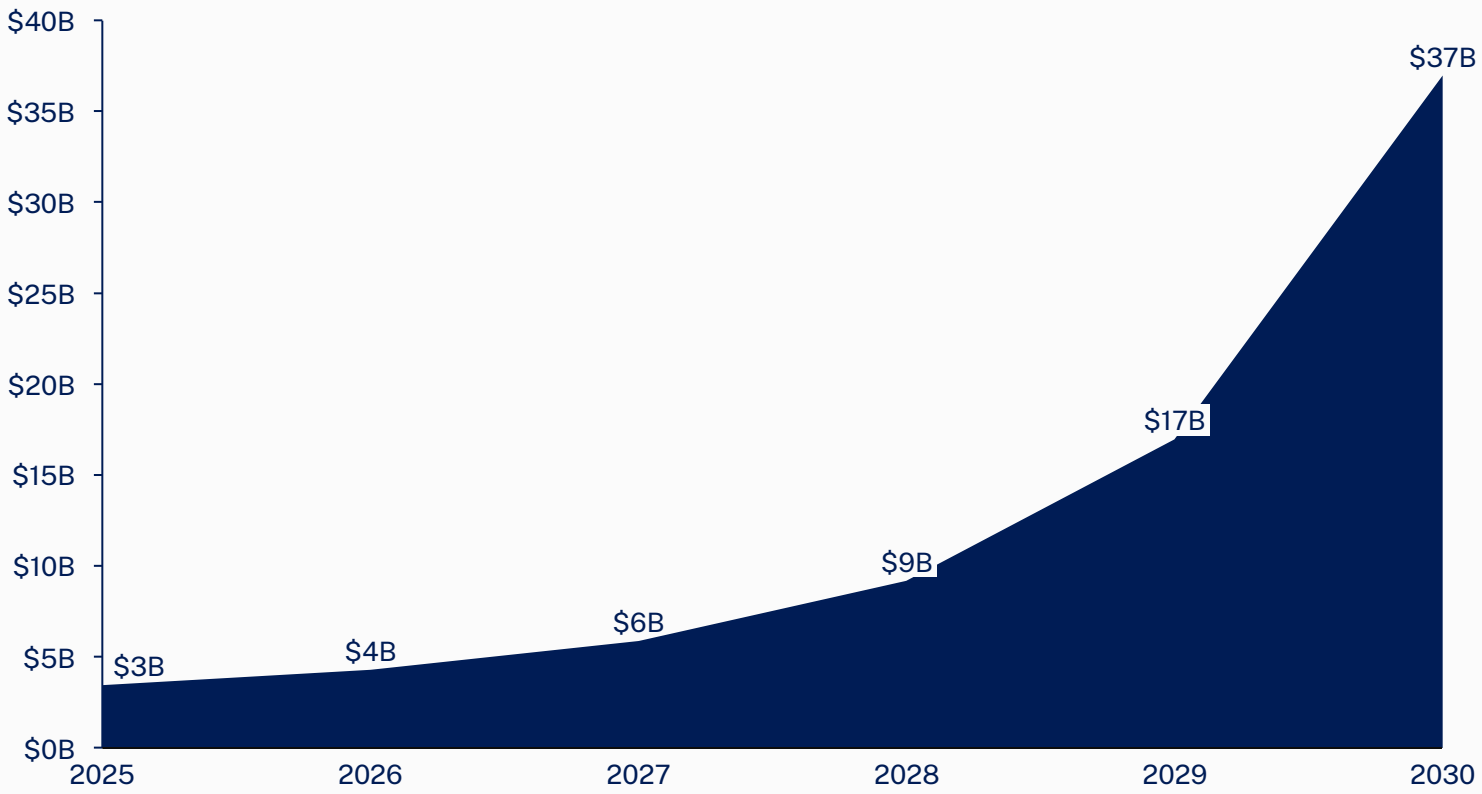
That headline number hides who bears the brunt of these losses. By 2030, the damage from a single loss incident is wildly uneven across business sizes: on average an incident will lose a large business around $4.5 million, versus about $38,000 for a micro business, a gap of more than 100 times. The number of loss incidents a year is also varied, with just 3,000 incidents across the largest businesses, while the smallest face nearly 250,000 incidents.

Among large businesses, those holding sensitive data and facing heavy regulation are predicted to take the worst hit: healthcare companies face losses of about $7.2 million per incident while financial companies face about $5.4 million in losses. This is due to the data they hold, the rules they answer to, and the cost of being knocked offline.

Source: Mandala analysis.

Australian business cyber incident losses

$, 2025-2030



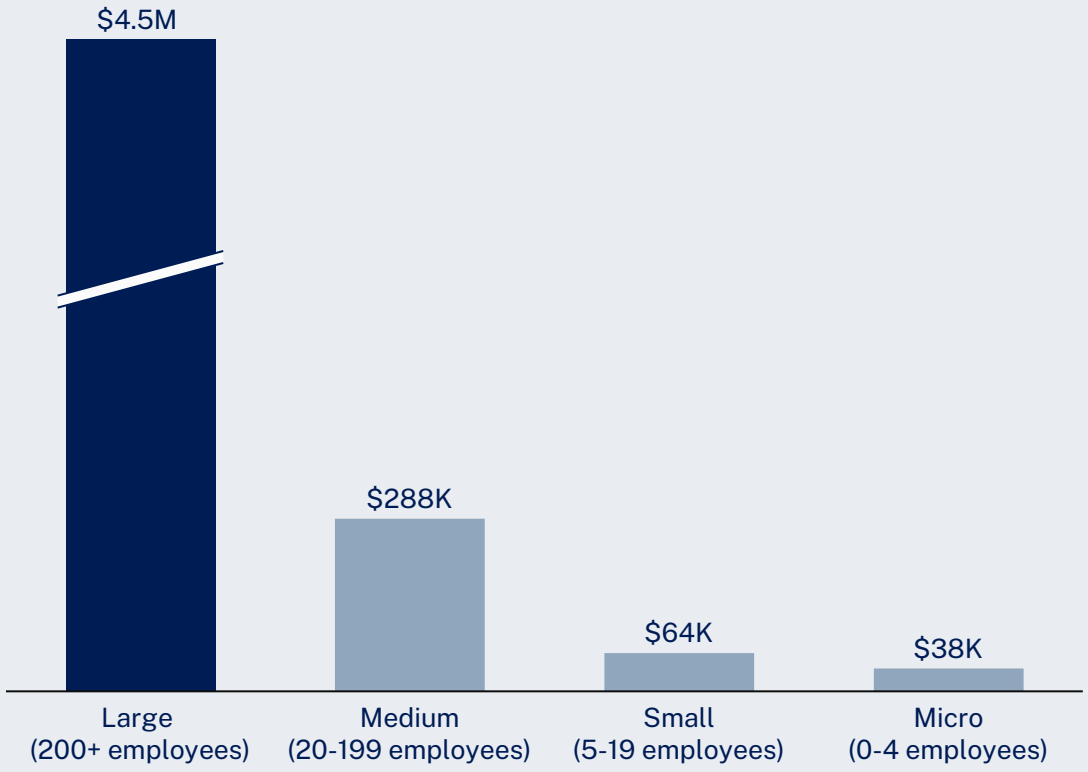
Improvements in AI capability drive increases in the frequency and severity of cyber attacks. Without defensive investment beyond current levels, Australian businesses could lose up to **$37B** annually to cyber incidents by 2030, more than **1.3 per cent** of GDP today.

Source: Mandala analysis.

Large businesses face the greatest exposure to cyber incident losses...

Losses per incident by business size in 2030

$, 2030



Sources: IBM (2025) *Cost of a Data Breach Report;* APRA (2025) *National Claims and Policies Database Statistics;* Australian Cyber Security Centre (2025) *Annual Cyber Threat Report 2024–2025;* Mandala analysis.

...with businesses operating in regulated industries at greatest risk

Top industries by large business loss per incident in 2030

$, 2030

The healthcare industry is predicted to see the highest losses per incident, at **$7.2M**. This is driven by the sensitive data, regulatory obligations, and losses from operational disruptions.

Note: Energy companies extract, produce or distribute fuel and power (ANZSIC Divisions B, C & D). Financial companies manage money, lending and insurance (Division K). Healthcare companies provide medical treatment and care (Division Q). Manufacturing companies manufacture physical goods and materials (Division C). Technology companies provide computing and internet services (Divisions J & M).

Sources: IBM (2025) *Cost of a Data Breach Report;* ABS (2013) *Australian and New Zealand Standard Industrial Classification (ANZSIC);* Mandala analysis.

By 2030, 76% of cyber losses will be locked in at the point of the incident

Of the $37 billion in projected losses, the majority is incurred when the incident occurs rather than in its aftermath. Two categories, business interruption and direct losses to threat actors, together account for $28 billion, or 76 per cent of the total.

Business interruption is the largest single component at $17 billion, nearly half of all losses. It reflects the cost of systems being taken offline, whether through malware, ransomware, denial-of-service attacks, or cascading failures at third-party suppliers. The organisation is not defrauded here; it is simply unable to operate.

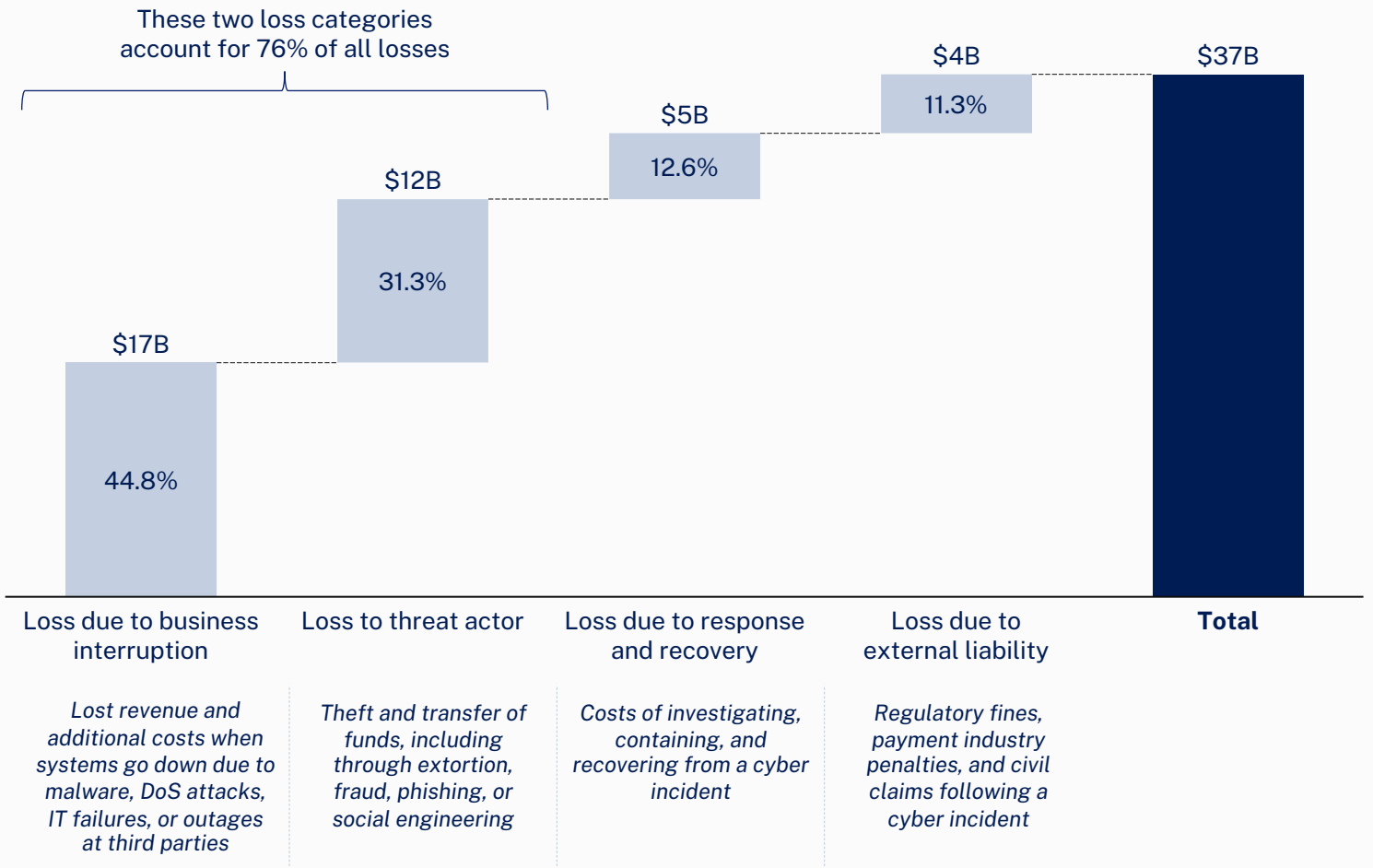
Direct loss to threat actors is the second largest category at $12 billion, close to a third of the total. These are the losses where funds or data leave the organisation entirely, through extortion, theft, fraud, phishing, and social engineering.

By contrast, the cost of responding and recovering is considerably smaller. Forensic investigation, crisis management, technical remediation, and the restoration of normal operations, together with the regulatory and legal consequences pursued by bodies such as APRA and the OAIC, account for less than a quarter of total losses combined. Prevention therefore carries substantially greater value than recovery. Most of the financial damage is determined at the moment of the incident, not during the response.

Note: Numbers may not sum due to rounding.
Source: Mandala analysis.

Cyber incident losses by outcome in 2030

$, 2030



Note: Numbers may not sum due to rounding.
Sources: Verizon Business (2025) *2025 Data Breach Investigations Report;* Mandala analysis.

Cyber incident losses can ripple throughout a supply chain, magnifying exposure for other businesses

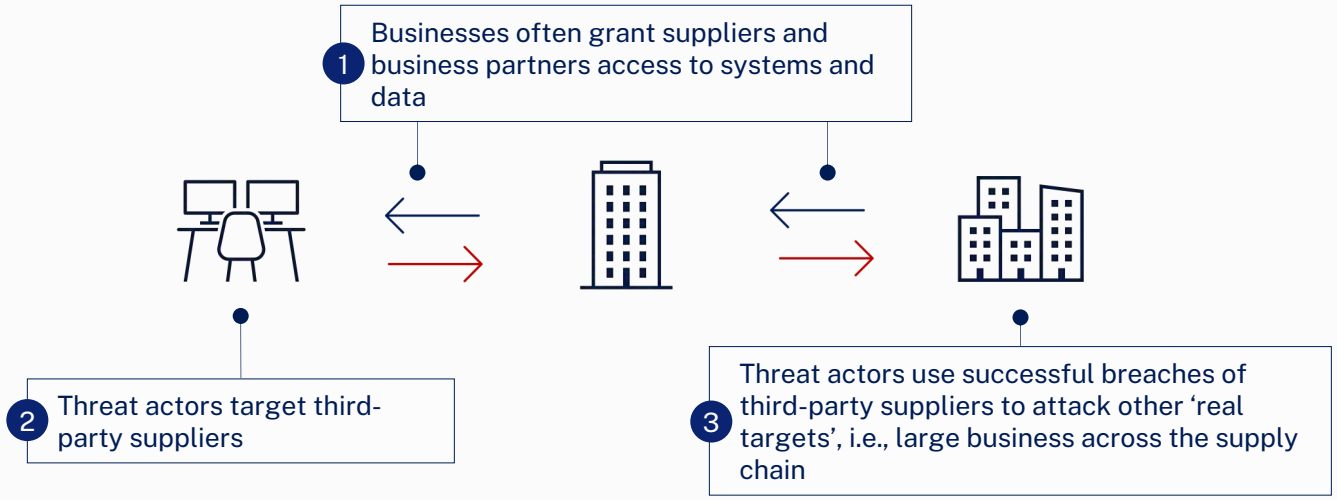
Third party suppliers are a common entry point for attackers on other businesses. Attackers target third-party suppliers as they have legitimate access to other downstream businesses, and some suppliers are small businesses with weaker security. Attackers can turn a compromise of a third-party supplier into a large-scale compromise against many businesses.

Attackers breach the supplier, steal its credentials, compromise its systems, and use that trusted access to reach into the partner networks behind it. Because one supplier often connects to many customers, a single breach could result in breaches of multiple large firms downstream at the same time. In the Canvas breach example, student and faculty data was compromised not through breaches at universities or schools directly, but through a learning management system that was used at those institutions.

The damage does not stop at the initial breach. The first hit is direct: attackers steal data and disrupt operations at the business itself. From there it can flow outward, turning into regulatory penalties, lost contracts and customers, reputational and brand damage, litigation, higher insurance premiums, and a falling share price, long after the intrusion is over. For universities, the fallout to the Canvas breach included disruptions to examination schedules, regulatory notification obligations, and exposure to litigation risk as data custodians.

Sources: Verizon Business (2025) *2025 Data Breach Investigations Report;* Australian Cyber Security Centre (2025) *Annual Cyber Threat Report 2024–2025*; Mandala analysis.

Ripple effects across supply chain cyber breaches



Case study: Canvas data breach, April-May 2026

- In April 2026, attackers gained unauthorised access to Canvas, a widely used university learning platform. A ransomware attack followed weeks later, locking students and staff out of study materials during exam period.
- Instructure reached an agreement with the attackers to secure the data, and engaged a consultant to conduct forensic analysis and provide monitoring.
- The company now faces legal action in California over the exposure of personal information.

Sources: Instructure (n.d.) *Security Incident Update & FAQs*; NBC 7 San Diego (2026) *Lawsuit filed in San Diego vs. Canvas developer following cyberattack*; CNN (2026) *What we know about the Canvas hack that has impacted thousands of schools*; Mandala analysis.

A ransomware attack on a technology provider led to disruptions throughout European airports

On 19 September 2025, a ransomware attack on a shared technology provider disrupted passenger processing services at airports that included London Heathrow, Brussels, Berlin, Dublin and Cork. The software was used by dozens of airlines across several airports, so the compromise of a single shared technology provider led to disruption across the European aviation network. AI will increase the likelihood of cyber attacks, including similar ransomware attacks.

Airlines were forced to switch to manual check-in of passengers and alternative baggage-processing arrangements, causing delays and cancellations. By the fourth day, roughly 90% of flights at Heathrow, 88% at Brussels, and 94% at Berlin Brandenburg were delayed, with average delays ranging from about 30 minutes at Heathrow to around an hour at Berlin Brandenburg. Disruptions to check-in and baggage systems at Berlin Brandenburg and Brussels persisted a week after the incident. An internal memo also indicated that over 1,000 computers at Heathrow may have been affected, requiring in-person rather than remote fixes. Affected airports and airlines were liable for compensation claims, rebooking, extra staffing, and lost passenger goodwill.

Flight delays of 30 minutes to 1 hour at airports across Europe

More than 1000 computers affected and needing fixing at Heathrow

Source: BBC (2025) *EU cyber agency says airport software held to ransom by criminals*; TechCrunch (2025) *European airports still dealing with disruptions days after ransomware attack*; Al Jazeera (2025) *Cyberattack on European airports caused by ransomware, EU finds*; EuroNews (2025) *European airports face another day of disruption on Monday after weekend cyberattack*; Mandala analysis

European airport outage

September 2025

Compromised open source packages can impact a large number of organisations

In early 2026, TeamPCP, a financially motivated cyber threat actor group, carried out a campaign in which it compromised at least four widely used open source software packages to introduce malicious code. The group poisoned trusted software packages used by developers, allowing the activity to extend to downstream organisations using the affected software.

The attackers injected malicious code into automated runtimes and package registries on open-source collaboration platforms. The malware extracted sensitive data such as secret access tokens and Secure Shell (SSH) keys. These stolen credentials unlocked access to further packages and repositories. Downstream victims included an AI recruitment firm, the European Commission's cloud infrastructure, and OpenAI.

Australian investigation teams assessed that TeamPCP activity affected more than 1,000 organisations worldwide, exposed more than 500,000 credentials and resulted in the theft of 300 GB of data. Remediation costs were assessed as potentially reaching hundreds of millions of dollars.

300GB of stolen data across more than **1,000 organisations**

These types of cyber attacks are likely to be exacerbated by AI

Sources: Wired (2026) *A Hacker Group Is Poisoning Open Source Code at an Unprecedented Scale*; Bitdefender (2026) *Shai-Hulud hackers: two men charged over TeamPCP's global supply chain crime spree that hit OpenAI, and thousands more*; Unit 42 (2026) *Weaponizing the Protectors: TeamPCP's Multi-Stage Supply Chain Attack on Security Infrastructure*; Mandala analysis

Open source code poisoning

February-March 2026

1 AI is changing what an attacker can do, making cyber incidents both more frequent and more severe

2 Australian businesses could face losses reaching $37 billion a year by 2030 as AI-driven threats grow

3 Businesses can avoid most of the losses by investing in foundational cyber defences, saving the economy $24 billion a year

4 Appendix

Businesses can take preventative and responsive actions to improve their cyber security

Businesses can take preventative actions before a cyber incident and responsive actions after a cyber security incident to minimise losses.

Preventative actions include establishing a cyber security strategy, assessing risks and aligning to a risk management framework, and taking actions to minimise those risks. A cyber security strategy should be consistent with the objectives of a business, and it is a key foundation for other actions that businesses can take for ensuring a secure by design approach. Assessing cyber security risks enables businesses to prioritise protective actions, which can include managing identities and access control with multi-factor authentication, providing training for employees, and securing data by ensuring an appropriate data management and disposal policy to reduce exposure.

Responsive actions include identifying and containing an incident quickly through appropriate monitoring and logging, taking actions to mitigate it, communicating with government, stakeholders, and customers, and rapidly restoring operations afterwards. Businesses should continuously monitor for indicators of compromise and stolen credentials to identify issues quickly, and implement response plans to contain an incident such as through exercising plans for response, crisis communications, and business continuity. Determining the lessons learnt is also a critical part of recovery.

Many businesses have gaps in cyber security protections, including foundational measures such as password management and multi-factor authentication. Taking action is increasingly important given the changing nature of threats from AI.

Sources: NIST (2025) *Incident Response*; NIST (2020) NIST SP 800-53 Rev. 5; Mandala analysis.

Key business actions to improve cyber security

Prevent and reduce the likelihood of cyber incidents

Before an incident

- **Govern:** Establish a strategy for managing cyber security risk and ensure it is followed
- **Identify:** Understand cyber security risks to a business, and prioritise opportunities for improvement
- **Protect:** Take actions to minimise cyber security risks

Cyber security incident

Detect and respond to reduce the impact of cyber incidents

During and after an incident

- **Detect:** Rapidly find and analyse indicators of a potential cyber security incident
- **Respond:** Take actions to mitigate and analyse a cyber incident, and to provide communication and reporting
- **Recover:** Restore operations and provide communication to reduce effects after a cyber security incident

Sources: NIST (2024) *The NIST Cybersecurity Framework (CSF) 2.0*; NIST (2025) *Cybersecurity Framework Profile for Artificial Intelligence (Cyber AI Profile): NIST Community Profile*; Mandala analysis.

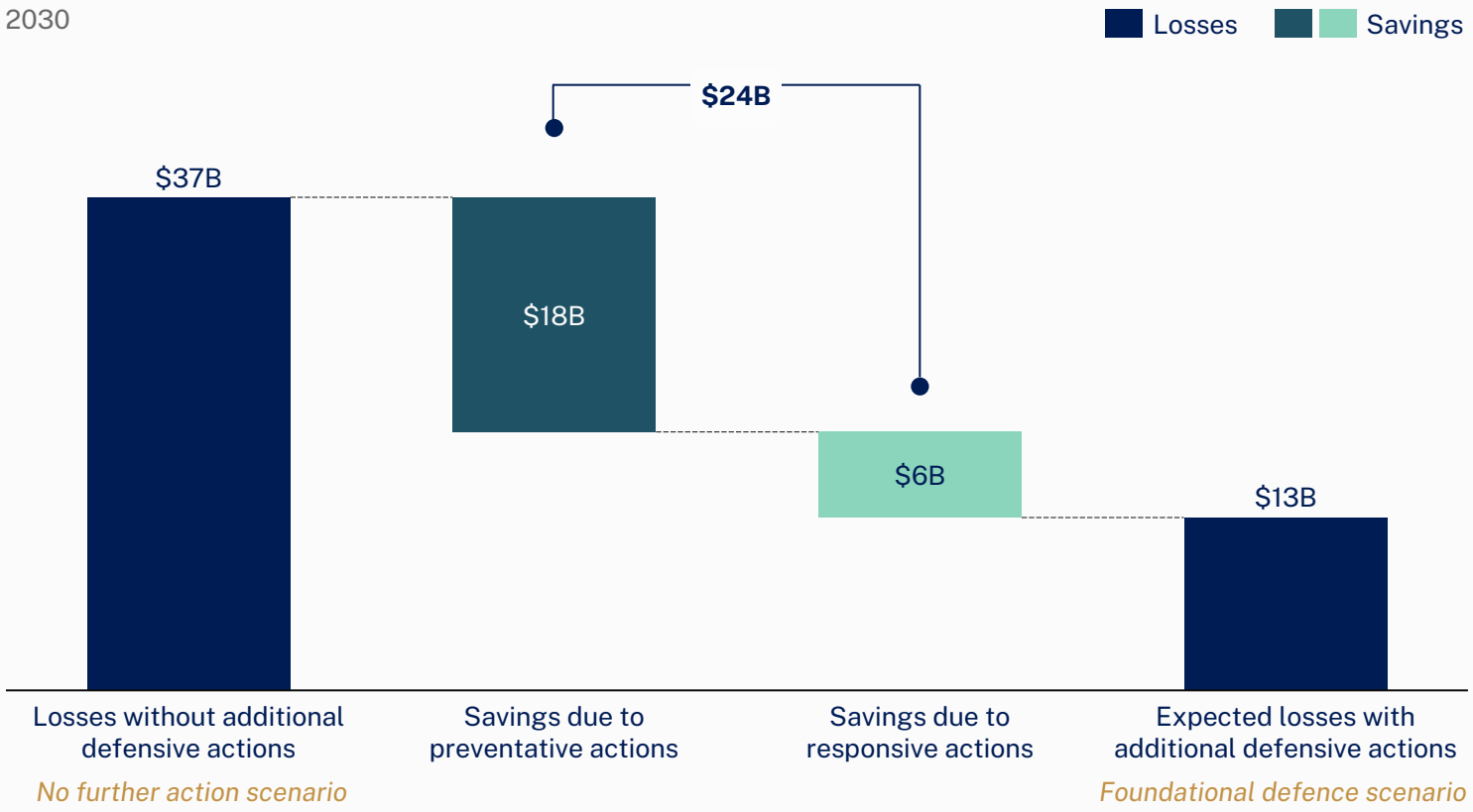
Taking action can reduce cyber incident losses by $24 billion in 2030 primarily due to preventative actions

Australian businesses are projected to lose $37 billion to cyber incidents in 2030 under the no further action scenario. This scenario holds cyber security measures at current levels while the threat from AI-enabled attacks grows between 2025 and 2030.

Under the foundational defence scenario, adopting a foundational set of defensive actions that are available to most businesses today could reduce losses by $24 billion, meaning they face losses of $13 billion. Preventative actions account for $18 billion of the reduction, including enforcing multi-factor authentication, patching known vulnerabilities, and securing AI deployments. Responsive actions account for a further $6 billion, mainly through continuous monitoring that detects intrusions earlier.

Losses of $13 billion could still occur, for three main reasons. **First**, foundational controls can struggle to keep pace with zero day vulnerabilities identified by AI. **Second**, AI has reduced the time to exploit a vulnerability, meaning attackers can operate faster than controls can remediate. **Third**, AI increases the scale and sophistication of attacks and challenges existing defences through accelerating reconnaissance, writing malware and creating increasingly convincing social engineering lures.

Annual reduction in cyber incident losses by taking additional foundational defensive measures



Foundational defensive actions cut expected losses by around two-thirds, mostly through preventative measures. Losses of $13B could still occur in the foundational defence scenario since attacks using AI can increase speed and scale, surpassing the ability of humans to respond.

Sources: NIST (2024) The NIST Cybersecurity Framework (CSF) 2.0; NIST (2026) *CSFv2.0-to-SP-800-53-Rev-5-2-0*; MITRE Centre for Threat-Informed Defense (n.d.) *NIST 800-53*; MITRE Centre for Threat-Informed Defense (2024) *Project Summary;* Mandala analysis.

Australian businesses should uplift their cyber security investments in line with global industry standards

Cyber security is currently underfunded across much of the economy. Among CEOs of organisations that are lacking cyber resilience, inadequate funds is cited as the single greatest barrier to increasing resilience.[1] In a separate survey, 90 per cent of respondents said their organisation needs a moderate-to-high level of investment to protect against cyber incidents, including AI-driven threats.[2] Investment in prevention in particular needs to be considered in comparison to the cost of an attack. In the future, funding drivers may include legislative requirements to manage cyber risk.

Australian businesses in total spent 4.4 per cent of their IT budgets on cyber security in 2025, in comparison to the global industry standard of around 7-10 per cent.[3,4,5] Businesses in Australia are spending less than global peers, as US and EU businesses are estimated to have spent 10.9 per cent and 9 per cent of their 2025 IT budgets on cyber security, respectively.[6,7]

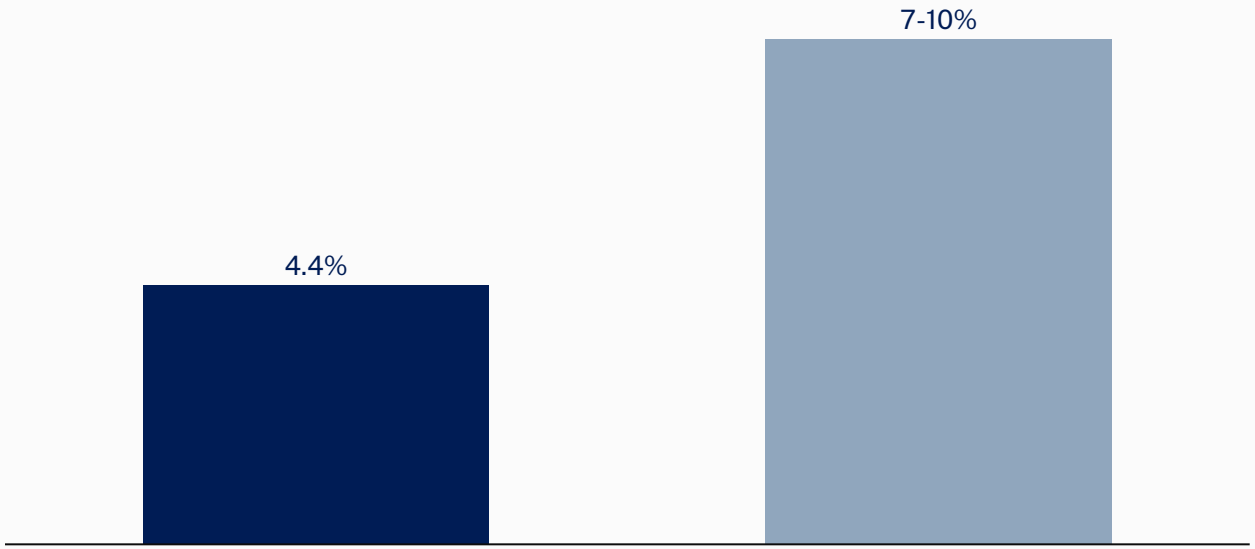
For an Australian business with $5 billion in annual revenue and $500 million in annual IT spend, 4.4 per cent of the IT budget would result in allocating $22 million to cyber security. Meeting global industry standards of 7-10 per cent of the IT budget means allocating $35-50 million to cyber security.

Sources: 1 WEF (2026) *Global Cybersecurity Outlook 2026*; 2 Allianz (2026) *Allianz Risk Barometer*; 3 Gartner (2026) *Gartner Forecasts Information Security Spending in Australia to Reach Over $7.5 Billion in 2026*; 4 Gartner (2025) *Gartner Forecasts IT Spending in Australia to Exceed $172 Billion in 2026*; 5 CSO Online (2019) *How much should you spend on security?*; 6 IANS & Artico (2025) *2025 Security Budget Benchmark Report*; 7 ENISA (2025) *NIS Investments 2025*; Mandala analysis.

Cyber security investments as a proportion of IT budgets

%

Australian businesses on average spend around 4.4 per cent of their IT budgets on cyber security, lower than global industry standards of around 7-10 per cent.



	Australian businesses	Global industry standard
*Implied **cyber spending** for a large business with $500 million in IT spend*	*$22 million*	*$35-50 million*

Notes: Australian businesses spent $6.9 billion on information security in 2025, estimated to be 4.4 per cent of $158 billion in total IT spending.
Sources: Gartner (2026) *Gartner Forecasts Information Security Spending in Australia to Reach Over $7.5 Billion in 2026*; Gartner (2025) *Gartner Forecasts IT Spending in Australia to Exceed $172 Billion in 2026*; CSO Online (2019) *How much should you spend on security?*; Mandala analysis.

To enable robust cyber resilience, businesses should improve their cyber defences via preventative and responsive capabilities

KEY ACTIONS		FOUNDATIONAL ACTIONS: BUILD RESILIENCE TO THWART THREATS	FOUNDATIONAL ACTIONS: IMPROVE AI IMPLEMENTATION	ADVANCED ACTIONS: USE AI FOR CYBER DEFENCE
Prevent and reduce the likelihood of cyber incidents	**Govern**	▪ Fund modernisation of legacy architecture ▪ Enable secure-by-design by implementing enterprise wide SDLC	▪ Create and maintain AI governance at executive, operational, and technical levels ▪ Align AI plans to risk posture and privacy risks ▪ Create solution architecture and AI use cases	▪ Support analysis and documentation for cyber security governance objectives
	Identify	▪ Assign the continuous monitoring of assets to accountable owners ▪ Hold third party providers accountable to enterprise level cyber risk tolerance thresholds	▪ Identify and create controls for sanctioned AI use cases ▪ Identify unsanctioned AI use and redirect users to sanctioned use cases	▪ Accelerate processes for monitoring and logging risks and vulnerabilities ▪ Determine the relevant controls given threat intelligence
	Protect	▪ Implement identity governance, privilege access management and non-human identity key management ▪ Make multi-factor authentication mandatory and use passkeys or ephemeral credentials ▪ Shift static patching cycles to continuous threat exposure management ▪ Institute a defined AI-SDLC process	▪ Implement secure-by-design AI architecture with measures such as AI workload sandboxing, human in the loop controls, and data and agentic memory storage monitoring ▪ Isolate and monitor agentic attack surface artefacts such as skills files, MCP servers, and memory files	▪ Continuously monitor operating environments ▪ Utilising AI pattern matching capabilities to detect deviations ▪ Use agentic scanning capabilities to analyse code weaknesses or vulnerabilities, develop patches, test patches, and queue for deployment with human team approval
Detect and respond to reduce the impact of cyber incidents	**Detect**	▪ Define and implement a detection and response operating model or SOC ▪ Consume and integrate threat intelligence into control uplift	▪ Monitor AI and agentic system activities ▪ Deploy AI-specific gateways, and manage agentic actions	▪ Build custom agentic flows to assist with hypothesis-based threat hunts ▪ Assist with high noise signals for insider risk
	Respond	▪ Implement endpoint and network detection and response capabilities ▪ Identify and retain an incident response service provider	▪ Accelerate specific automation or security event monitoring activities ▪ Scan and categorise unsanctioned agentic usage	▪ Invest in people, skills, and processes to manage complex AI systems that may provide security event capabilities
	Recover	▪ Create and practise system level recovery and incident response ▪ Create, maintain and test immutable backups of critical data	▪ Analyse and create system-level recovery plans for agentic models and data ▪ Practise backup and recovery procedures where a critical service is reliant on AI	▪ Assist with post-incident reporting and communications ▪ Verify integrity of backups and system recovery objectives regularly

Sources: OpenAI (2026) *A call for collective action on cyber defense*; NIST (2025) *Cybersecurity Framework Profile for Artificial Intelligence (Cyber AI Profile): NIST Community Profile*; Mandala analysis.

Australia's cyber policy is shifting, but government can tackle the advanced threat from AI across four fronts

Australia's cyber policy is already shifting in response to the threat environment, but AI will widen the gaps the current agenda leaves open. Government has a window to move from ambition to delivery across four fronts.

The Australian Cyber Security Strategy entered its second phase, Horizon 2 (2026 to 2028), in June 2026, with lifting the cyber maturity of small and medium businesses among its objectives. Acting on these planned programs is important given the enhanced vulnerability of smaller firms.

The independent review of the Security of Critical Infrastructure Act 2018 (SOCI Act) was released in March 2026, and two-way information sharing between critical infrastructure operators and government remains unresolved. Determining how mandatory threat intelligence sharing will work is a key step in strengthening national cyber capability.

Beyond these, government should build the future cyber workforce, and give business practical direction on AI: both on the risks of using it and on harnessing it, including frontier and open weight models, for defence.

The recommended actions opposite set out how government can act across each of these fronts.

Sources: Department of Home Affairs (2026) *Horizon 2: Expanding our reach (2026–2028)*; CyberCX (2025) *Consultation on developing Horizon 2 of the 2023–2030 Australian Cyber Security Strategy*; Dr Jill Slay AM (2026) *Independent Review of the Security of Critical Infrastructure Act 2018*; Department of Home Affairs (2026) *Streamlining and modernising the Security of Critical Infrastructure Act 2018*; Mandala analysis.

Recommended actions for government

Facilitate threat intelligence sharing
Collaborate with industry to establish a model for real-time threat intelligence sharing across critical infrastructure sectors, with a goal to formalise these arrangements through legislation

Develop the cyber workforce
Work with industry to provide clarity on skills and labour market forecasting so educational institutions understand the **cyber skills demand**, develop **cyber education programs for school students**, and facilitate **employer placement schemes for tertiary students**

Provide sector-specific AI cyber security guidance
Clear sector-specific guidelines on the use of AI should be **updated regularly,** and **critical infrastructure sectors** should undergo **audits** to ensure they are following guidelines

Secure access to frontier AI and create standards for open weight models for cyber defence
Negotiate government-backed agreements with AI developers for Australian access to frontier models for defensive cyber operations, and consult with industry to develop guidelines on the use of open weight models for cyber defence

Sources: Department of Home Affairs (2026) *Horizon 2: Expanding our reach (2026–2028)*; CyberCX (2025) *Consultation on developing Horizon 2 of the 2023–2030 Australian Cyber Security Strategy*; CI ISAC Australia (2026) *The SOCI Act Review and the Case for Mandatory Intelligence Sharing*; Australian Signals Directorate (2026) *Careful adoption of agentic AI services*; OpenAI (2026) *A call for collective action on cyber defense*; Hugging Face (2026) *Be Ready Before the Attack: A Practical Guide to Self-Hosting an Open Model for Cyber Defense*; Mandala analysis.

1 AI is changing what an attacker can do, making cyber incidents both more frequent and more severe

2 Australian businesses could face losses reaching $37 billion a year by 2030 as AI-driven threats grow

3 Businesses can avoid most of the losses by investing in foundational cyber defences, saving the economy $24 billion a year

4 **Appendix**

Glossary

TERM	DEFINITION
Threat actor	A malicious individual or group that carries out cyber attacks against organisations or systems
Zero-day vulnerability	A software flaw unknown to the vendor that has not yet been patched, leaving systems exposed to exploitation
Ransomware	Malicious software that encrypts a victim's data and demands payment for its release
Phishing	A social engineering technique where attackers impersonate trusted entities to deceive targets into revealing credentials or sensitive information
Malware	Malicious software designed to disrupt, damage, or gain unauthorised access to systems
Deepfake	AI-generated synthetic audio, video, or images used to impersonate individuals
Multi-factor authentication	A security control requiring users to verify identity through two or more methods before gaining access
Supply chain attack	A cyber attack that targets a less-secure supplier or third party to gain access to the systems of a larger organisation
Denial-of-service (DoS) attack	An attack that floods systems with traffic or requests to make them unavailable to legitimate users
AI cyber capabilities	The ability of AI models to perform tasks associated with cyber attacks, measured by AISI by the complexity and duration of tasks a model can complete reliably
Sandbox	An isolated computing environment used to safely test software without risk of affecting external systems
Legacy architecture	Older IT systems or infrastructure that may lack modern security controls and are difficult to update or replace
Privilege escalation	A technique where an attacker gains higher-level permissions within a system than they were originally granted
Lateral movement	The process by which an attacker moves through a network after gaining initial access, seeking higher-value targets or data
Exfiltration	The unauthorised transfer of data out of an organisation's systems

Sources: AISI (2026) *How fast is autonomous AI cyber capability advancing?*; Verizon Business (2025) *2025 Data Breach Investigations Report*; MITRE (2025) *Enterprise tactics;* Mandala analysis.

We examine the impact of changing AI cyber capabilities on cyber loss incidents

CHANGING CYBER CAPABILITIES OF AI MODELS

The modelling examines the effect of increases in the cyber capabilities of AI models on the frequency and severity of cyber loss incidents.

AI cyber capabilities:
- The annual growth of AI cyber capabilities from AI Safety Institute data[1] is compared to the growth in AI-driven cyber incidents from IBM data[2]
- The expected annual growth in AI cyber capabilities is translated into growth in AI-driven cyber incidents

Impact of AI on frequency of cyber loss incidents:
- The use of techniques for cyber loss incidents is based on MITRE CITD Sightings data[3]
- Differing annual growth rates are applied for techniques, based on Anthropic's assessment of the level of AI applicability[4]

Impact of AI on severity of cyber loss incidents:
- The increase in cost per cyber loss incident is determined using IBM data on the higher cost of incidents that use AI[2]

IMPACT ON CYBER LOSS INCIDENTS

The modelling focuses on the impact of AI on cyber loss incidents. Cyber loss incidents are a type of cyber security incident where a loss occurs, and cyber security incidents are a type of cyber security event that compromises business operations.

Key definitions:
- **Cyber security event:** An occurrence of a system, service, or network state indicating a possible breach of security policy, failure of safeguards, or a previously unknown situation that may be relevant to security.[5]
- **Cyber security incident:** An unwanted or unexpected cyber security event that has either compromised business operations or has a significant probability of compromising business operations.[5]
- **Cyber loss incident:** A cyber security incident that has caused a financial loss[6]

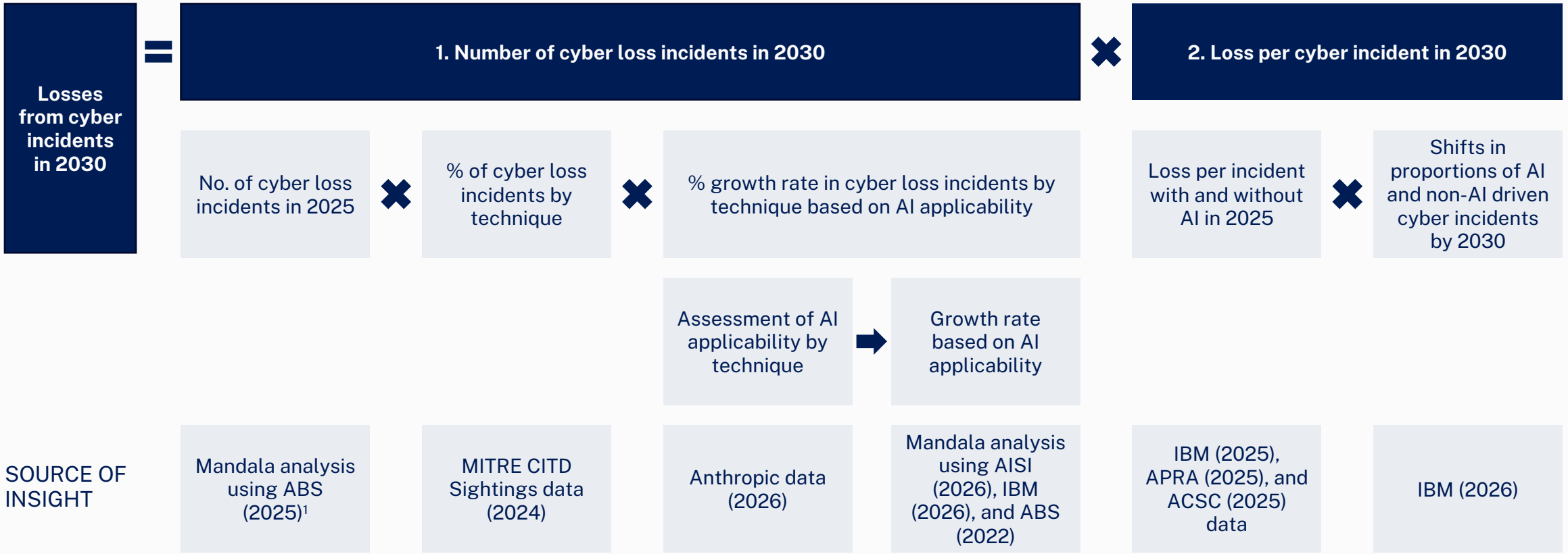
The number of Australian cyber loss incidents is determined using ABS data.[7,8]

Scenarios modelled

The modelling estimates losses under two scenarios, reflecting different levels of business response to the rising AI-driven threat.
- **No further action:** Cyber security measures are held at current levels while AI cyber capability continues to improve to 2030. This scenario isolates the effect of the changing threat environment and gives the upper estimate of $37 billion in annual losses.
- **Foundational defence:** Businesses increase investment in cyber defence and adopt additional foundational defensive measures available to most businesses today. Annual losses fall to $13 billion. Losses persist because attacks using AI can increase speed and scale, surpassing the ability of humans to respond.

Sources: 1 AISI (2026) *How fast is autonomous AI cyber capability advancing?*; 2 IBM (2025, 2026) *Cost of a Data Breach*; 3 MITRE Centre for Threat-Informed Defense (2024) *Project Summary;* 4 Anthropic (2026) *Mapping AI-enabled cyber threats: Insights from the LLM ATT&CK Navigator;* 5 Australian Signals Directorate (n.d.) *Glossary;* 6 Curti et al. (2023) *Cyber risk definition and classification for financial risk management;* 7 ABS (2022, 2025) *Counts of Australian Businesses, including Entries and Exits*; 8 ABS (2023, 2026) *Characteristics of Australian Businesses*; Mandala analysis.

Cyber losses in 2030 depend on the volume of cyber loss incidents and the losses per incident



Note: *All key assumptions, including those linking growth in AI cyber capability to the frequency and severity of cyber loss incidents, were reviewed and calibrated with senior cyber security practitioners at CyberCX. Loss estimates were separately cross-checked against independent sources including the UK Cyber Security Breaches Survey 2025 and the Verizon 2025 Data Breach Investigations Report.*

Notes: 1 Number of cyber loss incidents calculated based on the number of businesses by size, multiplied by the proportion of businesses that experienced a cyber incident by size, multiplied by the proportion of those that did experience a cyber incident which had a financial loss by size. This assumes businesses that experience cyber incidents, on average, experience one incident in a year.

Sources: ABS (2022, 2025) *Counts of Australian Businesses, including Entries and Exits*; ABS (2023, 2026) *Characteristics of Australian Businesses;* MITRE Centre for Threat-Informed Defense (2024) *Project Summary;* Anthropic (2026) *Mapping AI-enabled cyber threats: Insights from the LLM ATT&CK Navigator;* AISI (2026) *How fast is autonomous AI cyber capability advancing?*; IBM (2025, 2026) *Cost of a Data Breach Report;* APRA (2025) *National Claims and Policies Database Statistics;* Australian Cyber Security Centre (2025) *Annual Cyber Threat Report 2024–2025;* Mandala analysis.

AI applicability is assessed at the technique level

Key: AI applicability level based on % of observations

- High: >4%
- Medium: >2% to 4%
- Low: >0.05% to 2%
- None: 0% to 0.05%

Excerpt of AI applicability to top 15 most frequently used techniques in cyber incidents

Technique	Technique name	Proportion of observations of the used techniques by AI accounts	Assessment of AI applicability by technique based on proportion of observations
T1059	Command and Scripting Interpreter	4.5%	High AI applicability
T1027	Obfuscated Files or Information	4.3%	High AI applicability
T1105	Ingress Tool Transfer	0.4%	Low AI applicability
T1112	Modify Registry	0.4%	Low AI applicability
T1070	Indicator Removal	1.4%	Low AI applicability
T1204	User Execution	0.4%	Low AI applicability
T1564	Hide Artifacts	0.5%	Low AI applicability
T1055	Process Injection	4.3%	High AI applicability
T1003	OS Credential Dumping	0.4%	Low AI applicability
T1021	Remote Services	0.4%	Low AI applicability
T1486	Data Encrypted for Impact	0.2%	Low AI applicability
T1091	Replication Through Removable Media	0.0%	No AI applicability
T1082	System Information Discovery	1.2%	Low AI applicability
T1047	Windows Management Instrumentation	0.1%	Low AI applicability
T1562	Impair Defences	3.4%	Medium AI applicability
		Source: Anthropic data	Source: Mandala analysis

Key outputs:
357K cyber loss incidents in 2030 are based on the AI applicability analysis changing the growth rate in the use of techniques between 2025 to 2030

Notes: Technique frequency from MITRE CITD Sightings data: August 2021 to September 2023, AI applicability from Anthropic data: March 2025 to March 2026. Techniques based on MITRE Enterprise Techniques v18.

Sources: MITRE (n.d.) *Enterprise Techniques*; MITRE Centre for Threat-Informed Defense (2024) *Project Summary*; Anthropic (2026) *Mapping AI-enabled cyber threats: Insights from the LLM ATT&CK Navigator*; Mandala analysis.

The growth rate in technique use is applied based on levels of AI applicability

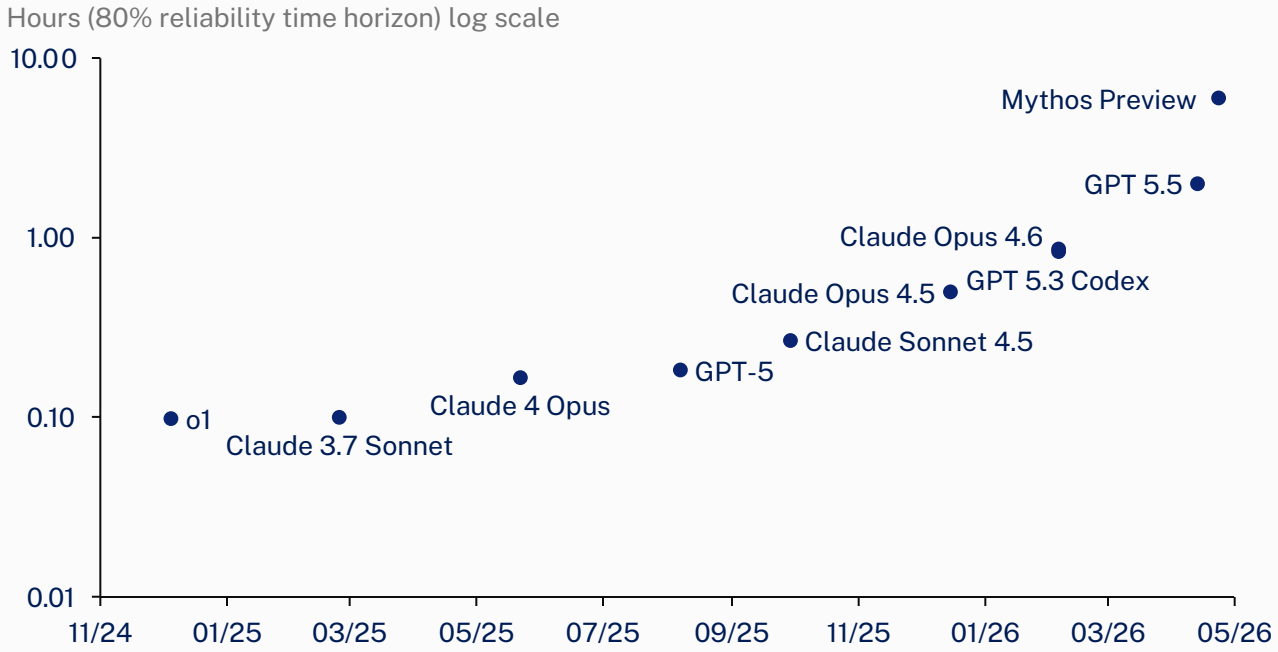
Annual growth rate in the use of techniques by level of AI applicability

AI applicability for technique	Description	2025-26	2026-27	2027-28	2028-29	2029-30
High	Annual growth starts at 56%, the observed rise in AI-driven incidents in 2025–26.[1] It increases linearly to 149% by 2029–30, in line with the projected pace of AI cyber capability (*see next slide for more details*).[2] Capability gains are assumed to take time to materialise in practice.	56%	79%	103%	126%	149%
Medium	56% rise in the number of AI-driven incidents over 2025-2026.[1]	56% per annum				
Low	Small expected change in the use of AI based on a 7% annual change in the number of Australian cyber loss incidents between 2022 and 2025.[3,4]	7% per annum				
None	No expected change in technique use based on AI.[5]	No growth				

Sources: 1 IBM (2026) *Cost of a Data Breach Report 2026*; 2 AISI (2026) *How fast is autonomous AI cyber capability advancing?*; 3 ABS (2022, 2025) *Counts of Australian Businesses, including Entries and Exits*; 4 ABS (2023, 2026) *Characteristics of Australian Businesses; 5* Mandala analysis.

The final year growth rate for highly AI-applicable techniques is based on linking incident volume to the rate at which AI cyber capability is improving

Trends in AI cyber capability[1]



- AISI measures **the length of cyber tasks, in human time, that a model can complete with 80 per cent reliability**. For example, Claude 3.7 Sonnet is expected to reliably complete 80 per cent of tasks that would take a human about 16 minutes to do.
- AISI's February 2026 re-estimate almost halves the time it takes an AI system to double its capabilities, from 8 months to 4.7 months. This corresponds to an increase in annual capability growth from 183 per cent to 487 per cent.

Translating capability into cyber incident volume growth by 2029-2030

1. Establish how cyber incident volume responds to capability

- AISI estimates a doubling of AI cyber capabilities every 8 months in 2025, which corresponds to AI cyber capability growing at 183% in the year to 2025[1]
- AI-driven cyber incidents rose 56% over 2025-2026[2]
- **Elasticity** = 56% ÷ 183% = **0.3**, which is held constant to 2030

2. Assess the pace of capability growth

- AISI estimates a doubling of AI cyber capabilities every 4.7 months in February 2026,[1] which corresponds to an **annual AI cyber capability growth** of **487%**

3. Apply the relationship to estimate growth in number of cyber incidents by 2030

- **487% * 0.3 = 149% annual growth in AI-driven cyber incidents**
- No. of cyber incidents assumed to reach this rate of growth in 2029 to 2030

Notes: Values may not add up due to rounding.
Sources: 1 AISI (2026) *How fast is autonomous AI cyber capability advancing?*; 2 IBM (2026) *Cost of a Data Breach Report 2026*; Mandala analysis.

The loss per attack estimation combines a 2025 baseline by business size with a severity uplift for AI-driven incidents

Loss per cyber incident in 2025

Business size	Estimated loss per incident in 2025 ($)	Source	Rationale
Micro (0-4 employees)	$33K	**Australian Cyber Security Centre** Annual Cyber Threat Report 2024–2025	Most representative source for small, micro, and sole trader businesses.
Small (5-19 employees)	$56K		
Medium (20-199 employees)	$252K	**APRA** National Claims and Policies Database Statistics	Australian-specific. Winsorised to remove outliers and exclude small and large businesses.
Large (200+ employees)	$3.9M	**IBM** Cost of a Data Breach Report	Reports Australian-specific figures oriented towards large businesses

Loss estimation for 2030

1. Set the baseline estimates

- **Loss per incident in 2025** is set separately for each business size, using the sources listed on the left

2. Split the baseline into AI and non-AI losses

- **AI-driven incidents** are assumed to cause **losses 1.2 times greater** than incidents that do not use AI (IBM 2026). Combined with the 2025 attack mix of 16% AI-driven, this gives loss per incident for each type separately.

3. Calculate the weighted average loss in 2030

- The AI and non-AI incident frequency mix in 2030 (from previous section) is used to calculate weighted average loss

These estimates were tested against sources like the UK Cyber Security Breaches Survey 2025 and Verizon 2025 Data Breach Investigations Report and reviewed with industry experts.

Sources: Australian Cyber Security Centre (2025) *Annual Cyber Threat Report 2024–2025;* APRA (2025) *National Claims and Policies Database Statistics;* IBM (2025, 2026) *Cost of a Data Breach Report;* Verizon Business (2025) *2025 Data Breach Investigations Report;* UK Government (2025) *Cyber Security Breaches Survey 2025;* Mandala analysis.

The foundational defence scenario maps defensive measures to techniques via NIST controls, lowering their AI applicability

Key: AI applicability level based on % of observations	Key: AI applicability level after mitigations
High: >4%	Medium
Medium: >2% to 4%	Low
Low: >0.05% to 2%	Low
None: 0% to 0.05%	None

Excerpt of mitigation impact on AI applicability for affected techniques

Technique impacted	Technique name	NIST CSF Functions	AI applicability before mitigation	AI applicability after mitigation
T1059	Command and Scripting Interpreter	Govern, Identify, Protect, Detect	High AI applicability	Medium AI applicability
T1005	Data from Local System	Identify, Protect, Detect	Medium AI applicability	Low AI applicability
T1190	Exploit Public-Facing Application	Govern, Identify, Protect, Detect	Medium AI applicability	Low AI applicability
T1578	Modify Cloud Compute Infrastructure	Govern, Identify, Protect, Detect	No AI applicability	No AI applicability
T1485	Data Destruction	Identify, Protect, Detect, Recover	Low AI applicability	Low AI applicability

How mitigations map to techniques

- NIST SP 800-53 controls are defensive measures that businesses can take to improve cyber security. These defensive measures are mapped to the relevant MITRE ATT&CK techniques that they help respond to or prevent.
- The NIST controls map to NIST CSF functions, that help identify the broad category of actions that businesses can take. Preventative savings are those from techniques covered by the Govern, Identify and Protect functions; responsive savings are those covered by Detect, Respond and Recover.

How the reduction is modelled

- Mitigated techniques are assumed to fall one level of AI applicability, reflecting reduced scope for AI to be used effectively against a defended technique. Techniques already assessed as having no applicability are unchanged.
- The lower applicability level attracts a lower annual growth rate in technique use (see slide 29), reducing the number of cyber loss incidents in 2030.
- The scenario assumes adoption of the foundational defensive set across all businesses.

Notes: NIST CSF functions are mapped to NIST 800 capabilities based on NIST's framework, and then NIST 800 capabilities map to ATT&CK techniques based on MITRE's framework.
Source: NIST (2026) CSF*v2.0-to-SP-800-53-Rev-5-2-0*; CITD (n.d.) *NIST 800-53*; MITRE (n.d.) *Enterprise Techniques*; MITRE Centre for Threat-Informed Defense (2024) *Project Summary*; Anthropic (2026) *Mapping AI-enabled cyber threats: Insights from the LLM ATT&CK Navigator;* Mandala analysis.

CyberCX
Part of Accenture

MANDALA