



Frontier AI cyber threat considerations for boards of directors

First published: August 2026

Introduction

The Australian Signals Directorate (ASD) prevents, disrupts and responds to attacks against Australian organisations every day. Understanding and managing cyber security risks, as with other business risks, is a key responsibility in protecting your organisation, shareholders and customers.

The Australian Institute of Company Directors (AICD) mission is to be the independent and trusted voice of governance, building the capability of a community of leaders for the benefit of Australian society. In recent years, the AICD has had a significant focus on educating boards of directors (boards) on better practice cyber security, artificial intelligence (AI) and data governance through guidance and education activities.

What boards need to know about frontier AI threats

Cyber threats are rapidly evolving due to the constant advances of frontier AI models. Frontier AI models are typically characterised by advanced reasoning, software development capabilities, natural language processing or multimodal capabilities, and are trained at scale using state-of-the-art techniques. Such AI models can perform complex tasks that exceed the capabilities of conventional automated tools. Frontier AI models have been widely available since early 2023 and continue to rapidly evolve.

Frontier AI models present threats that have implications for cyber security operations. This includes their ability to:

- identify vulnerabilities and rapidly weaponise them
- chain together multiple low severity vulnerabilities into high-impact compromises
- perform malicious cyber activities with little to no human oversight.

Newer frontier AI models are also dramatically increasing the speed, scale and ease at which vulnerabilities and weaknesses in systems and cyber supply chains can be identified and exploited, including by malicious actors that previously lacked the knowledge or skills to perform such activities. Combined with malicious actors exploiting weaknesses in proprietary frontier AI model guardrails and illegally distilling such AI models, these developments enable broader access to advanced capabilities and amplify traditional attacker-defender asymmetries. This may rapidly invalidate organisations' current risk tolerance. Overall, frontier AI models are currently, and will continue to, fundamentally transform both offensive and defensive cyber capabilities.

Boards should also be conscious that the implications of frontier AI models extend beyond their effect on cyber security operations. For example, boards should consider cyber supply chain risks associated with their reliance on particular AI vendors and the foreign ownership, control or influence risks this may pose.

These dynamics have clear implications for how boards should oversee proactive steps to enhance cyber security and resilience within their organisations. Organisations that fail to act now may be forced to accept greater business risk and exposure to losses from frontier AI threats.

What threshold questions should boards ask?

To assist with addressing current and emerging frontier AI threats, boards should be discussing the following threshold questions with management:

- How vulnerable is our organisation to AI-enabled attacks?
 - What assumptions underpin our current risk assessments, and how might frontier AI threats invalidate those assumptions?
 - Have we reviewed our risk tolerance in light of frontier AI threats and is that risk tolerance still appropriate?
 - If frontier AI models were used to identify and exploit weaknesses across our organisation, what areas of our business would be most exposed?
- Where could minor weaknesses in our systems and cyber supply chain be combined into a major cyber security incident?
 - Are there known weaknesses in our systems that could become materially more dangerous if identified and exploited at machine speed?
 - Are we relying on vendors and service providers without sufficient governance and oversight, including an understanding of their foreign ownership, control or influence?
 - Do we have visibility of the security and resilience posture of third and fourth-party suppliers within our cyber supply chain?
- Do we have control over our cyber security fundamentals, such as adherence with a recognised cyber security framework?
 - What legacy technology risks are we carrying and do we have a plan to remediate these risks?
 - Are we delaying addressing any weaknesses due to a perceived low exposure or severity?

- Would our cyber security operations remain effective if cyber attacks became more frequent, more targeted and more automated?
- Can we keep the business operational during a serious cyber security incident?
 - If one of our critical systems was compromised, could we continue operating our most important services?
 - If attacks moved from taking days to hours, could we still detect and respond to them effectively?
 - Have our incident response and business continuity plans been updated and tested to capture frontier AI threats?

What should boards focus on?

Frontier AI has the potential to significantly reduce vulnerability discovery and exploitation timelines from days to hours while significantly lowering the skill and knowledge barrier for malicious actors. Organisations should enhance their cyber security fundamentals now.

To give effect to a secure and resilient posture that is capable of responding to frontier AI threats, boards should oversee and challenge management on how their organisations are implementing the following strategies, spanning immediate priorities through to longer-term horizons.

Immediate priorities

- **Secure attack surfaces:** Systems are securely configured to approved and maintained baselines to reduce system exposure, including by reducing attack surfaces and attack paths, with configurations continually monitored and consistently enforced.
- **Reduce software vulnerabilities:** Vulnerabilities in systems are identified, documented, validated and prioritised for remediation or mitigation in a timely manner, with all remediation and mitigation actions verified for effectiveness.

Short-term priorities

- **Replace legacy systems:** Systems that are not capable of meeting cyber security requirements are managed using compensating controls, along with enhanced monitoring and assurance activities, to maintain an acceptable level of residual risk until they can be decommissioned or replaced.
- **Reinforce identity, credential and access management:** Robust and secure identity, credential and access management is used to establish, maintain and control access to systems and to support effective detection of identity and credential misuse.
- **Restrict unnecessary privileges:** Personnel and services, including AI agents, are granted the minimum access to systems required to undertake their duties.

- **Prepare for cyber security incidents:** Cyber security incident response, business continuity and disaster recovery plans for systems support continued business operations during cyber security incidents, and the resumption of normal business operations following cyber security incidents.

Medium-term priorities

- **Adopt AI for cyber defence purposes:** Artificial intelligence is used for cyber defence and is secure, controllable, human-supervised and used in an ethical and accountable manner.

Longer-term priorities

- **Modernise for the future:** Systems are planned, designed, developed, tested, deployed, maintained and decommissioned according to business criticality ratings and security and resilience requirements using Secure by Design and Secure by Default principles and practices.

What practical actions can boards oversee?

Below are practical actions recommended for organisations under each of the priority areas identified above. Boards are encouraged to engage with, and where appropriate challenge, management on their implementation as their organisation develops its response to frontier AI threats. This should include regular reporting and assurance.

Secure attack surfaces

- Defining approved configuration baselines that disable or remove unnecessary services, insecure settings, and unsupported or weak communication protocols.
- Deploying approved configuration baselines on systems, including infrastructure, operating systems and applications.
- Detecting and responding to configuration drift from approved configuration baselines through automation or continuous technical assessments.

Reduce software vulnerabilities

- Identifying known software and configuration weaknesses of assets through continuous and repeatable assessment activities, including vulnerability scanning activities.
- Applying security patches or updates within defined risk-based timeframes or, where operational impact is low, automatically.
- Remediating or mitigating vulnerabilities exceeding defined risk thresholds within defined risk-based timeframes based on their severity and exposure.

Replace legacy systems

- Establishing and maintaining a process for identifying, assessing and managing the cyber security risk associated with legacy systems until they can be removed or replaced.

Reinforce identity, credential and access management

- Identifying and disabling or removing default and unused accounts on a regular basis.
- Storing authentication keys, secrets and tokens in approved secure repositories with access controls, including multi-factor authentication.
- Regularly scanning codebases, configuration files, file shares and user collaboration platforms to identify unsecured keys, secrets and credentials.
- Using phishing-resistant multi-factor authentication for users authenticating to systems, including from untrusted, external or high-risk locations.
- Disabling legacy authentication protocols that bypass or weaken modern authentication controls.
- Ensuring that user, machine, system and service credentials are unique per system or account, particularly in high-privilege contexts.

Restrict unnecessary privileges

- Periodically reviewing and updating access privileges in alignment with current roles, duties and business requirements.
- Restricting permissions to create, modify or delete accounts, roles or access privileged functions or applications to authorised administrative roles.
- Provisioning user, administrative and non-person accounts with only the minimum access and privileges required to perform defined operational functions.

Prepare for cyber security incidents

- Reviewing cyber security incident response, business continuity and disaster recovery plans to ensure they remain fit for purpose.
- Regularly exercising cyber security incident response, business continuity and disaster recovery plans.

Adopt AI for cyber defence purposes

- Deploying suitable AI models for augmenting software development activities, such as identifying and remediating vulnerabilities and weaknesses for software before its release and periodically throughout its life.

- Deploying suitable AI models for augmenting security assessment activities, such as performing vulnerability scanning and vulnerability assessments.
- Deploying suitable AI models for augmenting security monitoring activities, such as identifying and triaging cyber security events.

Modernise for the future

- Designing systems to provide visibility, enforce authorised access, protect data, minimise required and default privileges, limit exposed services, and remove unnecessary functionality.
- Securing build and deployment processes to enforce controls consistently across systems.

Conclusion

Organisations that fail to respond to frontier AI threats leave themselves exposed. Boards should be pressing management to act now to ensure their organisations do not fall victim to current and emerging frontier AI threats, including by supporting targeted investments to enhance cyber security and resilience where appropriate.

Further information

For further information on the secure adoption and use of AI, consider the [suite of AI-related publications](#) ASD provides via their website.

The AICD has an [extensive set of digital governance publications](#) targeted at boards and directors, including [Data Governance Foundations for Boards](#), [A Director's Guide to AI Governance](#) and the [Cyber Security Governance Principles](#).

For more information on engaging with AI securely and responsibly, consider [Australia's Artificial Intelligence Ethics Principles](#). This is a voluntary framework produced by the Department of Industry, Science and Resources allowing for organisations to commit to ethical AI practices.

Contact details

If you have any questions regarding this guidance you can [write to us](#) or call us on 1300 CYBER1 (1300 292 371).

Disclaimer

The material in this guide is of a general nature and should not be regarded as legal advice or relied on for assistance in any particular circumstance or emergency situation. In any important matter, you should seek appropriate independent professional advice in relation to your own circumstances.

The Commonwealth accepts no responsibility or liability for any damage, loss or expense incurred as a result of the reliance on information contained in this guide.

Copyright

© Commonwealth of Australia 2026

With the exception of the Coat of Arms and where otherwise stated, all material presented in this publication is provided under a Creative Commons Attribution 4.0 International license (<https://creativecommons.org/licenses/by/4.0/>).

For the avoidance of doubt, this means this license only applies to material as set out in this document.



The details of the relevant license conditions are available on the Creative Commons website as is the full legal code for the CC BY 4.0 license (<https://creativecommons.org/licenses/by/4.0/legalcode.en>).

Use of the Coat of Arms

The terms under which the Coat of Arms can be used are detailed on the Department of the Prime Minister and Cabinet website (<https://www.pmc.gov.au/resources/commonwealth-coat-arms-information-and-guidelines>).



Australian Government
Australian Signals Directorate

ASD AUSTRALIAN
SIGNALS
DIRECTORATE
ACSC Australian
Cyber Security
Centre

**Australian Institute of
Company Directors**