agilient
Security Risk & Resilience

INFORM · ASSESS · PROTECT

SECURITY THREAT BRIEFING

The Malicious Use of AI

Artificial intelligence does not invent new threats so much as make familiar ones faster, cheaper and harder to spot. The same models that help defenders are being used to clone voices, build deepfakes, profile people and sites, and lower the skill needed to attack — across cyber, physical and personal security.

HOW THREAT ACTORS USE AI — THREE DOMAINS

Cyber

- Faster reconnaissance and vulnerability discovery
- Fluent, localised phishing and business email compromise
- Help to write and adapt malicious code
- "Dark" language models stripped of safeguards, sold to low-skill offenders

Physical

- AI-assisted hostile reconnaissance of sites and staff
- Aggregating public imagery, mapping and social data into a target profile
- Facial recognition to identify and track individuals
- Synthetic identities that pass rushed screening

Personal Information

- Voice cloning from a few seconds of audio
- Deepfake video for executive impersonation
- "Family in distress" and romance fraud at scale
- Identity theft and account takeover

WHO IS USING IT

Nation-states
Espionage and disruption; testing how far automation can run an operation.

Organised crime
Industrialised fraud, phishing and malware at volume.

Violent extremists
Propaganda, recruitment and operational planning.

Insiders / infiltrators
Fake identities and deepfakes to gain access.

Opportunists
Cheap, off-the-shelf voice and image tools.

BY THE NUMBERS

~US$25m
Lost by one firm after a deepfake video call impersonating its CFO and colleagues (Hong Kong)

US$893m
AI-related scam losses reported to the FBI in 2025, its first year breaking out AI

>$2bn
Scam losses reported in Australia in 2025, with AI-enabled deception a growing share

80–90%
Share of a 2025 state-linked cyber espionage campaign reportedly run by AI with limited human direction

WHAT SECURITY AND RESILIENCE TEAMS SHOULD DO

- ✓ **Treat AI as a cross-cutting threat** in an all-hazards security risk assessment, not only in the IT register.
- ✓ **Train people to expect pretexting** and deepfakes through security awareness training and clear escalation rules.
- ✓ **Strengthen workforce screening** against synthetic identities and remote-worker fraud.
- ✓ **Verify out of band.** Any request to move money, grant access or release data needs a check that does not rely on a voice or face alone.
- ✓ **Tighten reconnaissance exposure** — review what public data reveals about your sites, staff and routines.
- ✓ **Plan to keep operating** through an incident with tested business resilience and crisis arrangements.

Board question for tomorrow: **if someone cloned our chief executive's voice, which of our payment, access and data-release processes would catch it?**

agilient.com.au
SECURITY RISK ASSESSMENT · PROTECTIVE SECURITY · RESILIENCE

Sources: Anthropic (2025); Australian Signals Directorate, Annual Cyber Threat Report 2024–25; FBI Internet Crime Report 2025; CNN reporting on the Arup deepfake case; National Anti-Scam Centre. Figures point-in-time; verify before formal use.