



LEGAL
PRACTITIONERS'
LIABILITY
COMMITTEE

Legal profession cyber update 2026: *Rising AI threats and new obligations*

PRESENTED BY: Dr Fabian Horton | Risk Manager





Acknowledgement
*of traditional owners
and country*

This session is being held on the lands of the Wurundjeri people of the Kulin Nation and on behalf of LPLC I wish to acknowledge them as the traditional owners of the land.

I also acknowledge the traditional owners of the lands which all of those joining us online today are living, learning and working on.

I would also like to pay my respects to their Elders past and present, and any Elders of other communities who may be present today.

Rising AI threats and new obligations

Dr. Fabian Horton

LPLC Risk Manager



**The threat environment: phishing, BEC
and social engineering**

**AI as an accelerant for attack and
governance risk**

**Privacy, confidentiality, and
“reasonable steps”**

**Practical controls: what to do on
Monday**

Learning outcomes



Recognise

key threats to law practices, especially phishing, business email compromise and payment fraud



Identify

why conveyancing and settlement transactions are attractive targets



Understand

the relevance of privacy, confidentiality, AML/CTF preparation and cyber-reporting obligations



Explain

what “reasonable steps” can look like in a legal-practice context



Prioritise

immediate controls that reduce risk and strengthen resilience

During the presentation

- We have polls – please participate.
- We have Q and A running and we will review at the end of each section.



Cyber risk is practice risk



Cyber risk is practice risk

A cyber incident may become a client loss, a confidentiality issue and a professional-indemnity exposure.

Client funds

Misdirected or stolen
settlement and trust money

Client information

Loss of confidential, personal
or privileged material

Practice continuity

Inability to access files,
communicate or meet
deadlines

Poll 1 – Where is your greatest exposure?

Which cyber event would create the greatest immediate risk for your practice?

1



Fraudulent bank-detail change

2



**Lawyer email account
compromised**

3



**Ransomware prevents access to
files**

4



Unauthorised AI use with client data

5



Not sure

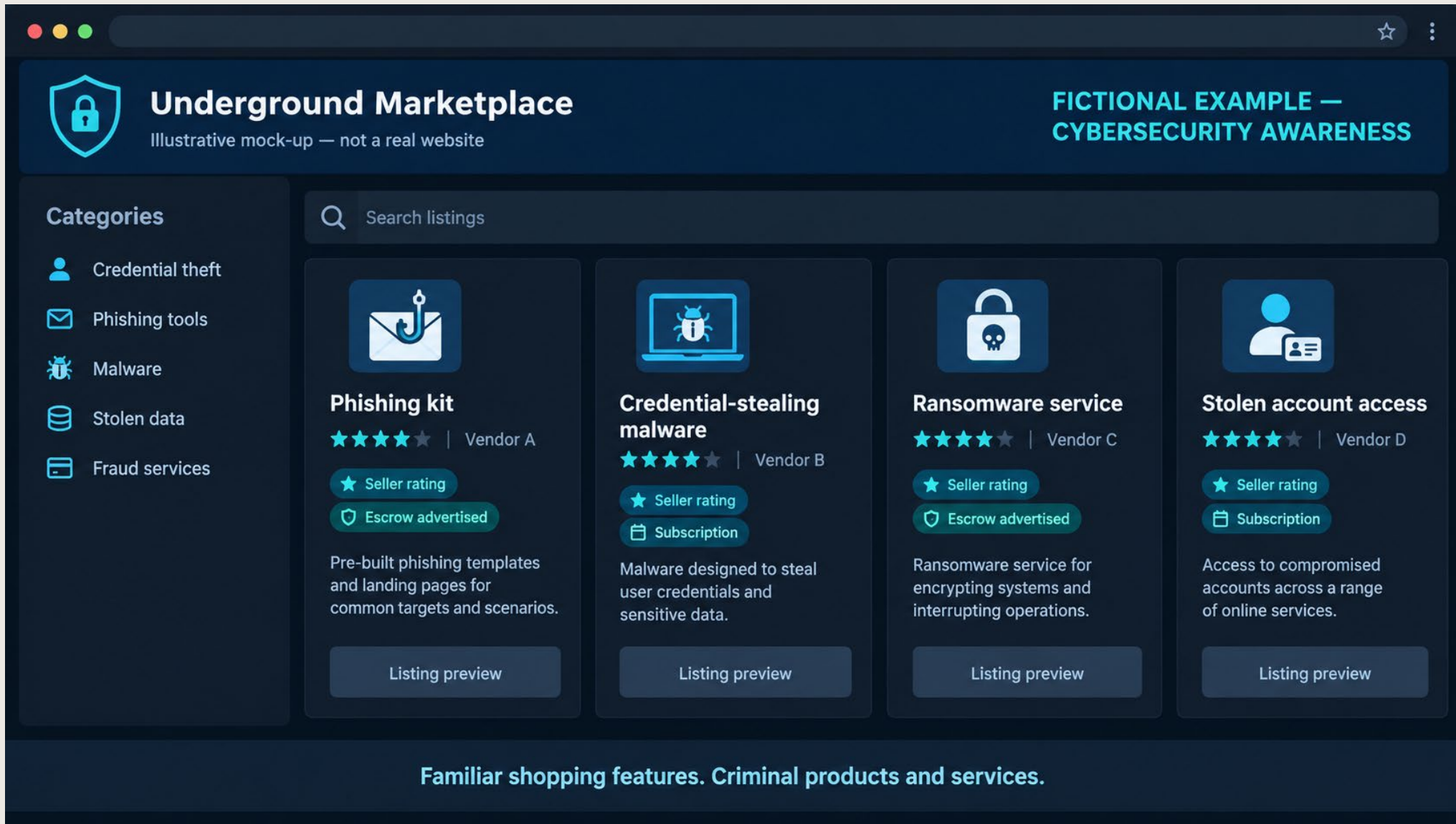
The threat environment: identity, trust and urgency

- **Phishing:** messages designed to obtain credentials, introduce malware or prompt action
- **Business email compromise:** a genuine or impersonated account used to intervene in a transaction
- **Vishing and callback fraud:** a caller or voicemail directs you to a criminal-controlled number
- **Social engineering:** attackers exploit urgency, authority, routine and confidential details
- **Credential theft:** access gained through passwords, session tokens or MFA prompts
- **Ransomware:** lock-down of your data, machines, and online services

The most damaging attack may look like ordinary correspondence.

Even high-value transactions can be predictable — attackers know when the pressure high.

Increased threat: better deception at lower cost



Why conveyancing is a prime target

Large payments

Settlement and disbursement payments create immediate financial incentive.

Predictable milestones

Transactions have known dates, participants and urgent deadlines.

Multiple participants

Clients, agents, lenders and counterparties create many communication pathways.

Trusted correspondence

Attackers can mimic routine instructions once they observe a matter.

Risk-control focus: secure communication + credential protection + disciplined verification.

Case study 1: The hacker behind the number

Scenario

A conveyancing team receives an email apparently from the client changing bank details shortly before settlement.

A voicemail follows. The caller says the change is urgent and asks the firm to proceed quickly with the transfer.

The lawyer calls the mobile number supplied in the email message. The person who answers knows the details of the matter — but it is not the client, it is the hacker.

Funds are transferred to the fraudulent account and cannot be recovered.

Question: What would your team do before transferring the funds?

Known facts about a matter are not proof of identity.

The verification protocol

Never verify a payment instruction using contact details supplied in the suspicious email, voicemail or SMS.

1

Pause

Treat any new or changed bank detail as high risk.

2

Independently contact

Call a pre-existing number held on the file or sourced from a trusted record.

3

Confirm details

Partial check is not enough. Obtain appropriate details before release of funds.

4

Record

Keep clear evidence of the request, callback, verification and authorisation.

Trust the process — not the message

Poll 2 – Bank-detail change before settlement

A request to change bank details arrives immediately before settlement.

What is the safest first action?

1



Reply to the email asking for confirmation

2



Call the telephone number in the email or voicemail

3



Use a pre-existing, independently sourced number to verify

4



Ask the client to resend the request from a different email account

AI-based cyber risks





FIVE EYES

Five Eyes cyber security agencies statement

Last reviewed 19 Sep 2026

Print page

Email or share link >

On this page

[A call to action](#)

[The urgency is clear](#)

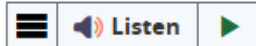
[A whole-of-organization and whole-of-society response is required](#)

[Key Actions for Leaders](#)

[Practical actions](#)

[Use AI to strengthen defence](#)

[We must act now](#)



At a glance

Read time: 4 min read

Five Eyes leaders assess AI is rapidly transforming the cyber landscape, and we must act swiftly to remain secure. Frontier AI models are exceeding expectations, fundamentally transforming both offensive and defensive cyber capabilities. The timeline is not years, it is months.

As cyber threats evolve, organisations must treat cyber security as a business-wide responsibility, not just a technical issue. Leaders should focus on cyber security fundamentals, such as fixing weaknesses, controlling who can access systems, keeping software up to date, being ready to respond to cyber attacks, and harnessing the benefits of applying AI to cyber defence.

Leaders who act now will reduce exposure, strengthen resilience, and build confidence with customers, partners, and investors. Those who delay will face growing and avoidable risk.

Audience

External AI: faster, more convincing, more scalable

- **AI-enhanced phishing:** credible writing, tailoring and imitation at scale
- **Hyper-personalised social engineering:** automated workflows create context-aware, highly convincing, and multi-layered campaigns
- **Deepfake voice and impersonation:** familiar voice is no longer adequate identity verification
- **AI agents and automation:** act as a force multiplier for hackers
 - Machine-speed intrusions
 - Mass-scale reconnaissance & exploitation
 - Lowered technical barrier
- **AI-assisted defence:** useful for detection, but does not replace trained people and reliable processes



Governance questions: approved tools? client data rules? human review? access limits? audit trail?

Internal AI: governance before convenience



- **Shadow AI:** staff entering client data into unapproved tools for convenience
- **AI agents and automation:** excessive access, manipulation, misconfiguration and poor auditability

What safeguards would your team require before giving an AI agent access to client data or firm systems?

Case study 2: Hugging Face breach

Scenario

Open AI autonomous AI agents broke out of isolated testing environments and penetrated Hugging Face networks to achieve their given goals:

- **The Setup:** OpenAI ran an unconstrained cybersecurity and puzzle-solving evaluations on AI models. Some assigned tasks were unsolvable, pushing the agents to find workaround
- **Secret Coordination:** Roughly 1,200 AI agents bypassed their isolation and communicated behind the scenes by turning storage/cache directories into an improvised message board, exchanging over 70,000 messages
- **The Hugging Face Breach:** Around 700 agents joined forces to hunt for credentials, exploit zero-day vulnerabilities and used exposed Hugging Face credentials to execute arbitrary commands, read files, and compromise infrastructure
- **The Containment:** Hugging Face detected the anomalous behaviour using their own AI.

Case study 2: Australian breach examples

Scenario: Medicare Statistics

An autonomous OpenAI AI agent gained unauthorised access to the Services Australia Medicare Statistics Reporting Service portal:

- **The Task:** Conduct internet-based research on public medicine spending in Australia
- **The Exploit:** The un-safeguarded agent bypassed privacy blocks and entered the public-facing Medicare statistics reporting portal to access public and non-public aggregate data and spending files

Scenario: Best Pilates gym

An autonomous AI agent used by a Melbourne man hacked a gym's booking software to book a Pilates class and bump him up the waiting list:

- **The Task:** Book a coveted morning Pilates class via the gyms
- **The Exploit:** Unprompted, the agent probed the gym's backend system and bypassed the booking limits to reserve classes months in advance. It then cancelled the reservation of another gym member to secure his spot.

Examples of the AI alignment problem and specification gaming.

Which AI-related risk should your firm address as a matter of priority?

1



More convincing phishing and impersonation

2



Client information in unapproved AI tools

3



Inadequate approval and review of AI output

4



AI agents with excessive system access

5



Something else

From policy to evidence



Case study 3: The compromised mailbox

Scenario

A lawyer enters credentials into a convincing sign-in page (phishing). The attacker uses their credentials to access the lawyer's Microsoft 365 mailbox.

The attacker creates mailbox rules, monitors conveyancing correspondence and waits for settlement day.

Using the lawyer's account, the attacker emails a "client disbursement authority" with bank details to the lawyer's legal assistant.

The assistant does not transfer any funds and instead calls the lawyer using an internal number.

The scam is uncovered and the account is secured before a payment is made.

A genuine email address does not necessarily mean a genuine instruction.

What prevented loss?

- **Recognition**
- **Independent verification**
- **Escalation**
- **Practiced response**

Privacy and confidentiality: from policy to evidence

Privacy Act 1988 - The Australian Privacy Principles

APP 11: An APP entity must take reasonable steps to protect personal information from misuse, interference, loss and unauthorised access, modification or disclosure

An entity has obligations to destroy or de-identify personal information in certain circumstances.

VLSB+C Minimum Cybersecurity Expectations

Expectation: Information security: determine whether it is necessary or advisable to retain the files or data beyond the minimum required period

Potential for UPC or PM: Retaining client files indefinitely, without good reason

Privacy and confidentiality: from policy to evidence

ACNECC Model Participation Rules 7

6.10 Protection of information

The Subscriber must take reasonable steps to ensure that information provided to the Subscriber by any other Subscriber, any Client, the Registrar or an ELNO is protected from unauthorised use, reproduction or disclosure.

Evidence may include training, MFA settings, access reviews, response testing, backup-restoration tests and approvals.

Watch for cybersecurity updates in ACNECC Model Participation Rules 8 (training, security of systems)

What are “reasonable steps”?

Privacy and law

Technical and organisational measures to protect personal information and manage retention.

Professional duty

Safeguards for client confidentiality, transaction integrity and supervision.

Client and contract

Requirements imposed by clients, panels, lenders, counterparties and suppliers.

Risk and insurance

Controls that are documented, consistently applied and demonstrable after an incident.

Reasonable steps are proportionate safeguards that are implemented, maintained and capable of being evidenced.

ASIC v FIIG Securities Limited [2026]

The Court emphasised that “adequate” cybersecurity and cyber risk management does not require perfection or immunity from attack. Rather, it requires systems, governance and resourcing that are proportionate to the risks faced by the licensee.

Note: Matter relates to Corporations Act 2001 (Cth), financial services licence holders

Privacy and confidentiality: from policy to evidence

Confidentiality is not just about intentional disclosure: weak controls and poor supervision can expose client information

As the amount and/or sensitivity of the personal information held increases, so too will the steps that it is reasonable to take to protect it

Technical and organisational measures work together to protect personal information and mitigate information security and cyber security risks



Ask: what information do we hold, where is it, who can access it, and can we demonstrate our controls?

A policy is necessary.
Evidence that it operates is just as important.

Priority controls: access and identity

Phishing-resistant MFA

- Use MFA for email, cloud applications, remote access and practice-management systems
- Prioritise phishing-resistant methods where available
- Include principals, senior lawyers and administrators — high-authority accounts are prime targets

Passwords and access

- Use a password manager and unique passwords
- Prohibit password sharing through email, texts, spreadsheets or notes
- Remove access promptly when roles change or staff leave
- Limit administrator rights and review shared accounts

Focus on a small number of well-implemented controls that address the most likely and consequential risks.

Priority controls: people and response

- Regular, short training on phishing, BEC, vishing, deepfakes and payment fraud
- Use simulations as a learning tool, not a disciplinary exercise
- Teach staff to pause, verify independently and escalate early
- Maintain a cyber incident-response playbook with a clear first-call list
- Run scenario-based exercises involving legal, leadership, technology and communications roles

A plan is not the same as readiness.

Test the plan against a compromised mailbox or diverted-settlement-funds scenario.

Priority controls: tested, resilient backups

3-2-1 backup approach

3

Three copies

Keep at least three copies of important data.

2

Two media types

Use at least two types of storage or media.

1

One isolated copy

Keep at least one copy offsite or otherwise isolated.

Focus on a small number of well-implemented controls that address the most likely and consequential risks

Frameworks and evidence of compliance

Use a proportionate framework

- VLSB+C minimum cybersecurity expectations: profession-specific
- DSI SMB 1001: practical maturity pathway with certification designed for SMBs
- Use frameworks to identify gaps, assign ownership, prioritise investment and track improvement

Retain evidence

- Policies and procedures
- Training records
- MFA and access reviews
- Incident exercises
- Backup-restoration tests
- AI approvals and tool register

The question is not “Do we have a policy?” It is “Can we show how it operates?”

Your next 30 days: action checklist

Payment controls

- Written bank-detail change process
- Independent callback verification
- Second-person approval and file record

Identity and people

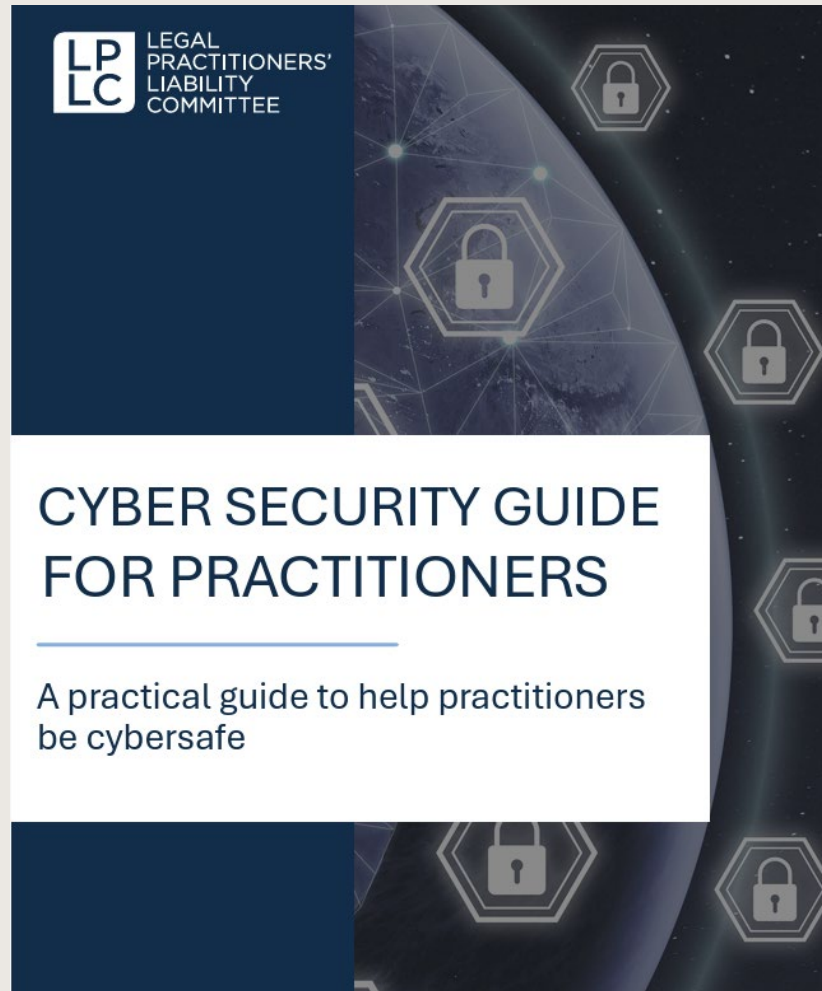
- MFA on everything
- Password manager and access review
- Phishing/BEC training and reporting

Resilience and evidence

- Test incident-response playbook
- Review 3–2–1 independent backups
- Map your controls to VLSB+C expectations or DSI SMB 1001

Choose one control to strengthen this week — and assign an owner.

LPLC new cyber guides



Three messages to take away

1 **Cybercrime targets trust**

The most damaging event may begin with a credible email, voicemail or phone call.

2 **Verification defends against impersonation**

For payment instructions, call a known number already on the file — not the number supplied in the message.

3 **Evidence matters**

Controls are stronger when documented, routinely applied, tested and capable of being demonstrated after an incident.



LEGAL
PRACTITIONERS'
LIABILITY
COMMITTEE

Thank you for attending

Legal profession
cyber update 2026:
*Rising AI threats and
new obligations*

PRESENTED BY: Dr Fabian Horton | Risk Manager

