Excite Cyber Whitepaper

Data Loss Prevention (DLP) as an Enabler for Secure AI Adoption

excite CYBER

1 - Summary

1.1 - Summary

Most conversations about AI start with capability and end with anxiety. The reality is more grounded, and more useful. AI has moved from experiment to operating model, and the organisations getting the most out of it are quietly reshaping how work gets done.

1.2 - The Value of AI

AI offers the chance to unlock high levels of immediate value. Teams are using Microsoft 365 Copilot to draft proposals in minutes instead of hours, summarise weeks of customer correspondence in a single page, and turn meeting transcripts into action lists before participants reach their next meeting. Engineering teams ship code faster with AI pair-programmers. Marketing, finance, and HR teams are automating the routine analysis and drafting that used to fill entire afternoons. These competitive advantages are being leveraged by Australian businesses of all sizes.

The next wave of innovation will be delivered through AI agents. Rather than just answering questions, agents take actions across multiple systems on your behalf. This may include raising support tickets, reconciling supplier invoices, triaging customer queries, scheduling work, even running parts of an onboarding process end-to-end. Platforms like Microsoft Copilot Studio and a growing ecosystem of low-code AI builders make this accessible without a software development team. For most leaders, the strategic question is no longer “should we adopt AI?” but “where can we get the most leverage from it?”

1.3 - Scale & Sovereignty

The answer to that question, almost without exception, comes back to data. Every AI tool, whether it’s Microsoft 365 Copilot, ChatGPT Enterprise, or a custom-built agent, works by surfacing, summarising, and reasoning over data. Mostly you and your organisation’s data. That makes data the single biggest variable in whether AI delivers value or causes problems. Well-organised, classified data produces useful, accurate output. Disorganised, unclassified data produces oversharing, surprises, and risk. The technology is the same; what differs is the foundation underneath.

1.4 - Impact on the workplace

AI usage is now mainstream & ubiquitous. McKinsey’s *2025 State of AI* found **88% of organisations regularly use AI** in at least one business function, and **79% are using generative AI** specifically. The Microsoft–LinkedIn 2024 *Work Trend Index* reports **75% of knowledge workers use AI** tools at work, and **78% of those bring their own**; this is what’s now widely called shadow AI.

IBM’s *2025 Cost of a Data Breach Report* puts numbers on the impact: 13% of organisations have already experienced an AI-related breach, and 97% of those lacked basic AI access controls. Breaches involving shadow AI cost an average of US$670,000 more than standard incidents. The common thread isn’t recklessness. Instead, it’s visibility the data and tools in use. Reco’s *2025 State of Shadow AI Report* found 86% of organisations don’t have a clear picture of how data flows to and from AI tools. You can’t protect what you can’t see.

1.5 - The pressures you’re already balancing

If you’re leading security or technology, you’re being pulled in several directions at once. The board wants visible AI adoption progress and the productivity gains other organisations are reporting. Audit and risk are asking sharper questions about data governance, the Australian Privacy Act reforms, and obligations under the Security of Critical Infrastructure (SOCI) Act. Cyber insurance underwriters are tightening requirements year on year. Enterprise customers are asking for evidence of data classification, encryption, and access controls before they sign. And underneath all of it, your security team is leaner than the threat landscape really warrants.

A coherent data security foundation answers more of these pressures simultaneously than almost any other investment. Knowing what data you have, classifying it, and applying consistent protection lets you say yes to AI with

confidence, makes audit and customer due diligence faster, supports your insurance position, and gives your security team meaningful signal to work with instead of noise.

1.6 - Where DLP fits — and why it's now an AI enabler

Data Loss Prevention used to be about stopping someone emailing a customer database to the wrong place. The discipline has matured. Modern DLP — built around discovery, classification, labelling, and policy enforcement — has become the practical foundation that makes secure AI adoption possible.

Microsoft Purview is the most accessible example for organisations already invested in Microsoft 365. Through a single platform you can:

- **Discover** sensitive data across Microsoft 365, endpoints, browsers, cloud apps, and structured data stores
- **Classify** content using built-in or custom sensitive information types and trainable classifiers
- **Label** files with sensitivity labels that travel with the content, wherever it goes
- **Protect** it with policies that determine what AI tools can read, summarise, or surface

When a user asks Copilot a question, Purview's labels and DLP policies are evaluated in real time. Confidential content stays confidential. Material that should be shared, is. The result is AI that's useful without being a liability.

1.7 - How Microsoft Purview organises, tracks, and secures your data

Purview is best understood as four connected capabilities operating across your data estate. None of them is new on its own — what's changed is how cleanly they now compose around AI use cases.

Discovery and classification scan your environment to identify where sensitive data lives. Purview's data map covers Microsoft 365 (SharePoint, OneDrive, Teams, Exchange), endpoints, and structured stores including Azure SQL, Synapse, Cosmos DB, and Amazon S3. Built-in sensitive information types detect common patterns — credit card numbers, tax file numbers, health identifiers — and trainable classifiers handle harder content like contracts, source code, or your own categories. Exact Data Match lets you classify against your own reference data, such as customer numbers from a CRM export.

Sensitivity labelling turns classification into a durable, organisational signal. Labels travel with the file across SharePoint, OneDrive, Teams, Outlook, Office apps, Loop, and Microsoft 365 Copilot, and can be auto-applied based on classification rules or selected by users. Each label can carry encryption, watermarks, headers, and external-sharing controls. The point is consistency: a "Confidential" file behaves the same way whether it's opened in Word on a managed laptop, forwarded as an email attachment, or surfaced in a Copilot response.

Data Loss Prevention enforces what should and shouldn't happen with that data. Modern DLP policies cover endpoints (including browser-based pasting into ChatGPT, Claude, or Gemini), email, Teams chats, SharePoint, and — most relevantly for AI — Copilot prompts and responses themselves. Microsoft's general availability of DLP for Microsoft 365 Copilot prompts means a policy can prevent confidential information from being processed by Copilot in real time, including for documents carrying specific sensitivity labels.

Insider Risk Management and audit close the loop with visibility. Activity Explorer and Data Security Posture Management for AI surface where sensitive data is being accessed by AI applications, by whom, and how often. Insider Risk Management correlates DLP signals with broader user activity to flag genuinely risky behaviour without drowning analysts in low-fidelity alerts. All AI prompt and response activity flows into the unified audit log for forensic review.

1.8 - Identity and zero trust: the bigger picture

Data protection and identity are inseparable. Purview's controls assume a working identity foundation — most commonly Microsoft Entra ID — and become significantly more powerful when paired with conditional access, privileged access management, and risk-based authentication.

This is also where data security becomes a meaningful contribution to a zero-trust strategy rather than a parallel programme. The three principles behind zero trust all touch the data layer directly:

- **Verify explicitly:** Conditional access policies in Entra can require stronger authentication when users access content carrying high-sensitivity labels or attempt risky actions on classified data
- **Use least-privilege access:** Sensitivity labels with encryption enforce permissions independently of file location, so even if a confidential file leaves an authorised location, it can't be opened without rights
- **Assume breach:** Audit logs, DSPM for AI, and Insider Risk signals give you data-layer telemetry to detect misuse early, which matters more now that AI-related breaches are taking longer to detect than traditional incidents

In practice, a piece of confidential customer data labelled in SharePoint is encrypted automatically, won't be returned in a Copilot response to a user without rights, will trigger an alert if anyone attempts to download it from an unmanaged device, and leaves a clean audit trail throughout. The capabilities aren't individually novel; the value is in how they now compose.

1.9 - A pragmatic, phased approach

There are key steps to move through in order to discover, organise and ultimately unleash greater value from your data sets.

1. **Discover:** The sensitive data you have, where it lives, who has access, and how it moves
2. **Classify and label:** The highest-priority data first: customer records, financial data, contracts, and IP
3. **Set proportionate policies: B**lock clear risks like pasting source code into public AI tools, and warn on grey areas
4. **Measure and refine:** Using Purview's built-in dashboards as your AI use scales

This is the same pattern whether you're rolling out Microsoft 365 Copilot to two hundred users or piloting a custom AI agent for customer service. The control plane is consistent; it's the data layer that needs the work.

1.10 - The bottom line

AI doesn't have to feel abstract or unmanageable. It works on your data, and once you know what data you have and how sensitive it is, decisions about AI become straightforward. The organisations that get this right won't necessarily be the ones with the most AI deployed. Instead, they'll be the ones whose data is in good enough shape for AI to actually be trusted with it. That's the work, and it's the most useful first step on any AI journey.

References

1. IBM. Cost of a Data Breach Report 2025, July 2025. newsroom.ibm.com

2. McKinsey & Company. The State of AI in 2025, 2025.

3. Microsoft & LinkedIn. 2024 Work Trend Index Annual Report, 2024.

4. Reco. 2025 State of Shadow AI Report, 2025.

5. Netskope. 2025 Cloud and Threat Report, 2025.

6. Microsoft Learn. Microsoft Purview data security and compliance protections for Microsoft 365 Copilot, learn.microsoft.com/purview, 2026.

7. Microsoft Learn. Sensitivity labels and Microsoft Purview