



Public Sector Innovation Week 2025

How government is
adopting AI for automatic
classification of data at
huge scale

SPEAKER INTRODUCTION

- CEO and co-founder of multi-award-winning solution Castlepoint Systems
- CIP, CISA, CISM, CDPSE
- Certified in project, change, and records management
- Cultural anthropology and linguistics background
- Actively involved in bringing ethical, global and sustainable practices to the sector

Australia's Most Outstanding Woman in IT Security and RegTech Female Entrepreneur of the Year. Australian Top 100 innovator.



About Us ▾

Resources ▾

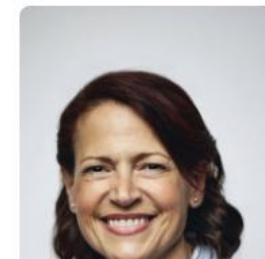
Solutions ▾

Platform ▾



Contact us

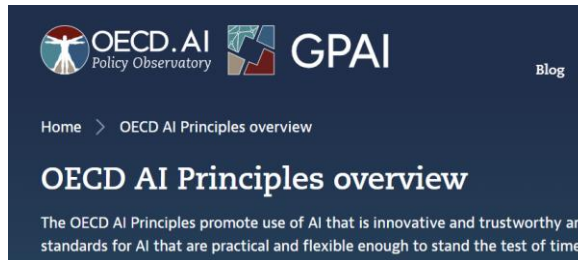
Control your risk, command your data



THE CURRENT RISK AND OPPORTUNITY



Ethics of Artificial Intelligence



Guidance

AI Playbook for the UK Government

The playbook offers guidance on using AI safely, effectively and securely for civil servants and people working in government organisations

From: [Government Digital Service](#)

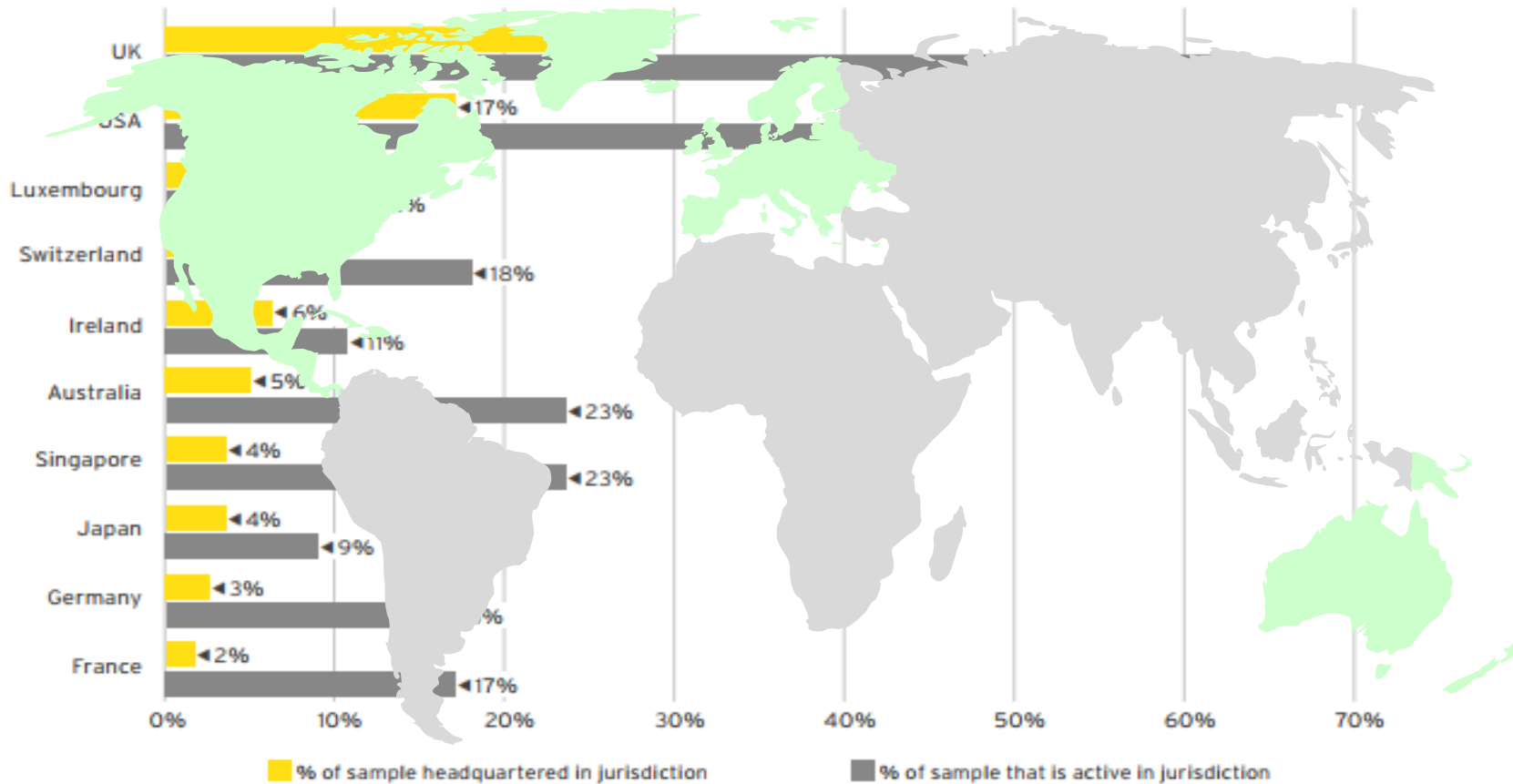
Published 10 February 2025

Explainability describes the ability to clarify how an AI system arrives at a given output or decision, such as explaining what factors lead to a loan application being granted or denied. Explainability may be impacted by the technologies used to build a system. You should consider this when designing your system.

Ensuring transparency and explainability can be challenging in the context of AI. Transparency can be limited by proprietary and 'black box' commercial tools, while explainability may not be possible for certain forms of machine learning, or may only be achievable at the cost of performance.

THE ROADMAP TO AUTOCLASSIFICATION ADOPTION

Figure E2: Top 10 RegTech markets, by % of firms present (headquarters or significant market share)



THE USE CASES FOR AUTOCLASSIFICATION

Autoclassification for text content has the following key use cases:

Discovery

Makes every asset discoverable based on its content and function. Provides contextual results, with risk and value specific to the organisation. Search indexes are created rapidly, without moving or changing content.

Cyber

Automatically detects sensitive and high-risk content, specific to the organisation. Shows the risk profile of content, allowing effective triage and impact management.

Privacy

Detects PII, PCI, and PHI, all identifying information, and complex sensitive content such as disability status, criminal history, or veteran status. Enables ESI eHold, and Data Minimisation per regulations.

Audit

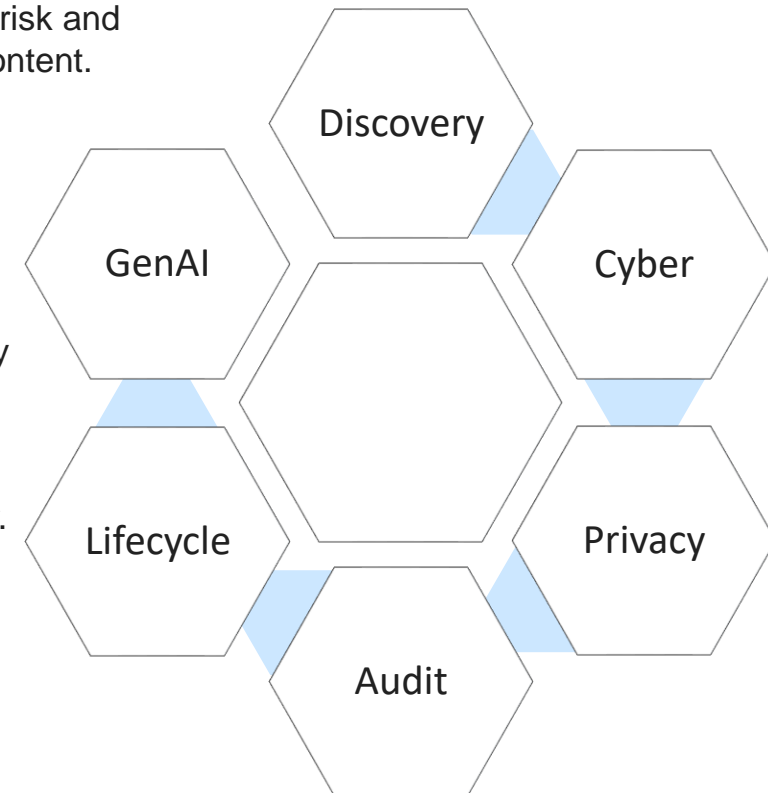
Logs, tracks, and alerts on events on data across the whole network, preserving events permanently.

Records Lifecycle

Automatically determines retention policy and disposes of records compliantly. Manages records compliantly 'in place' using XAI content-based autoclassification with very low overhead.

Generative AI Governance

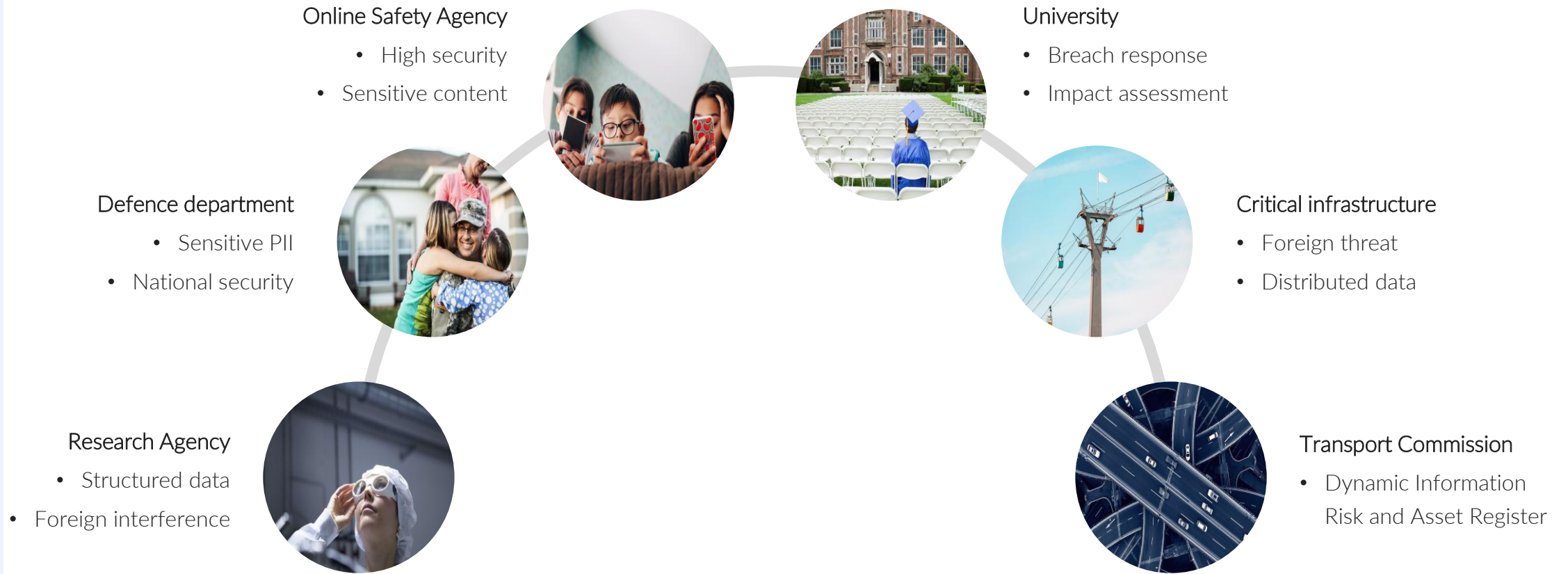
Enables quality and sensitivity oversight of GenAI user calls and system outputs, to help plan for and manage GenAI use.



THE CASE STUDIES – DISCOVERY



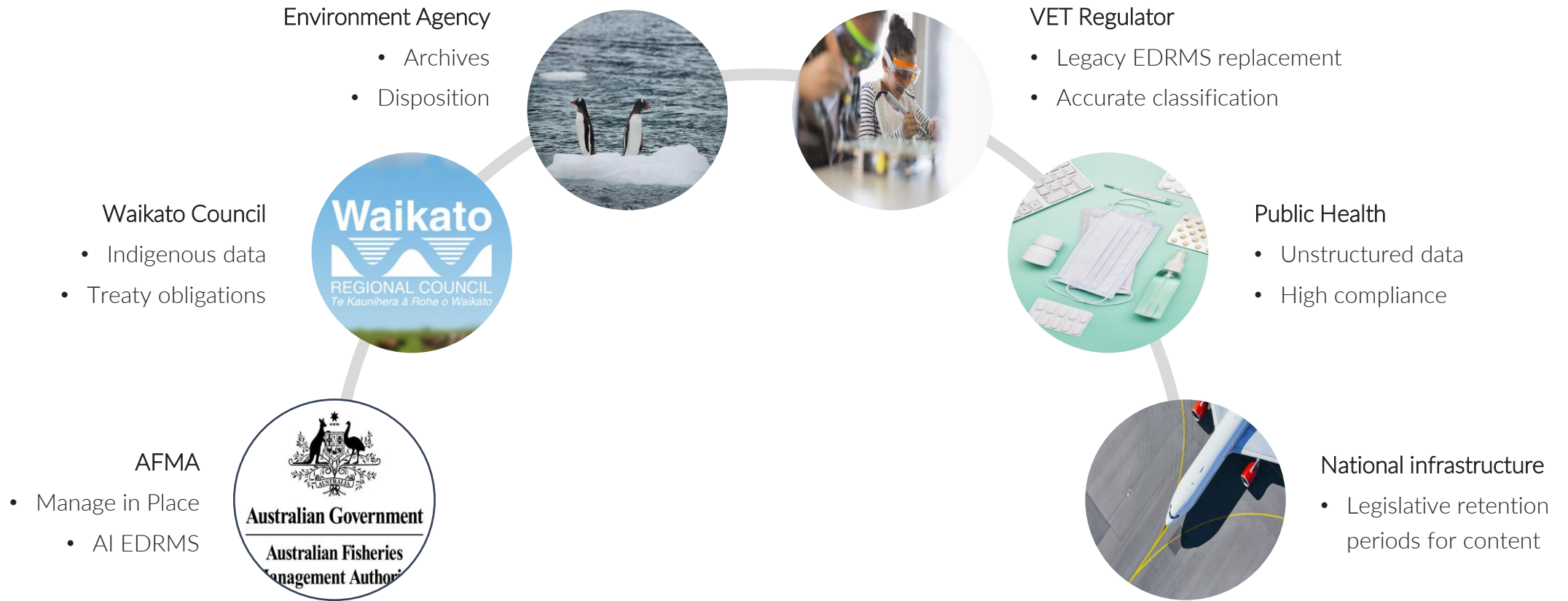
THE CASE STUDIES – SECURITY



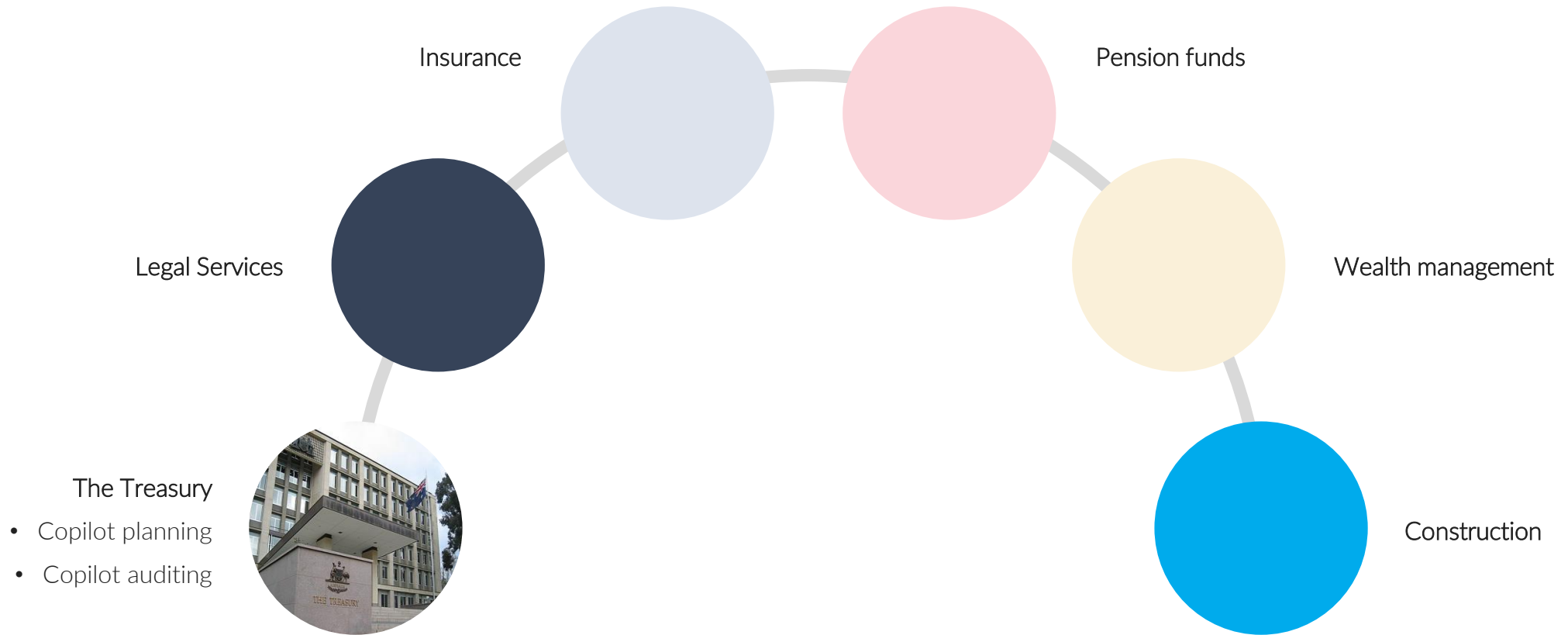
THE CASE STUDIES – PRIVACY AND DATA MINIMISATION



THE CASE STUDIES – RECORDS LIFECYCLE MANAGEMENT



THE CASE STUDIES – GEN AI GOVERNANCE



We **disable the labels**. Autoclassification based on content is more accurate, faster, and more scalable.

The old way

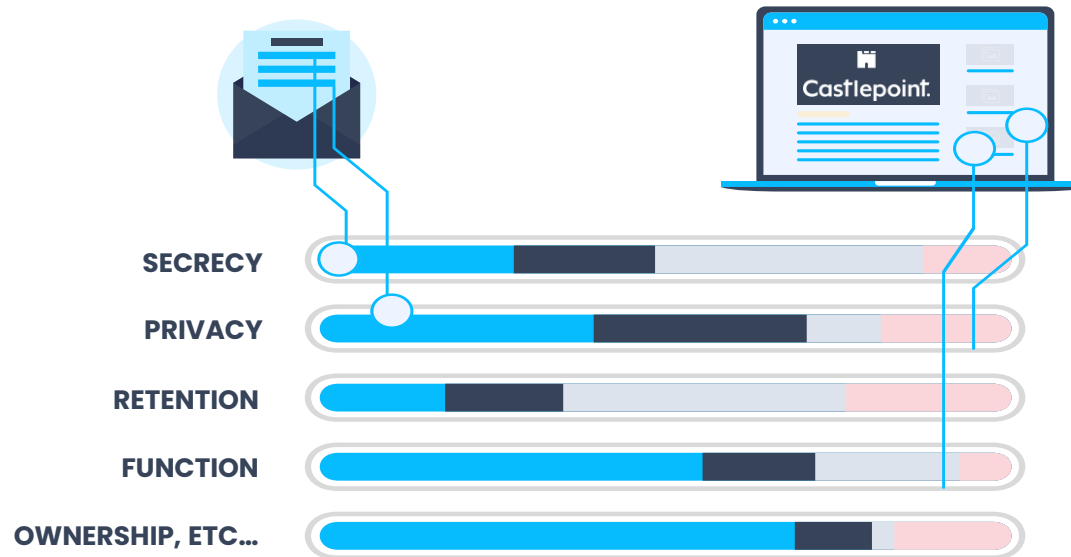
Every item must be tagged, labelled, named, or filed in a certain way to denote what it pertains to. But there are only so many labels that can be applied in practice, and only so much enthusiasm for staff to apply them (and maintain them accurately over time). Metadata fields are different in every system, and enterprise consistency is impossible.



- Limited available labels
- Manually added or mapped
- Inconsistent over time and systems

The Castlepoint innovation

Castlepoint reads all the content and metadata of your information objects, and determines what they are about. It captures all relevant labels for every item, matched against every rule, policy, regulation, and operational priority. But they are stored securely in the Castlepoint database, with no changes to the items.



- Unlimited labels
- Automatically captured
- Interactive through search, ontology, and classification interfaces
- Relates items across systems
- No impact on data integrity

Castlepoint turns any phrase or word in a document, email, web page, chat, or database entry into a label for that item, which can be used, alone or in any combination, for discovery, categorisation, and classification. No user input required, and no impact on source data or source systems.



contact: rachaelg@castlepoint.systems

