THALES
Building a future we can all trust

inception
Un evento de TrustDimension
EXECUTIVE SUMMIT
2026

Thales CSP: Helping you unlock the power of AI

Roman Baudrit

VP Sales, CSP LATAM

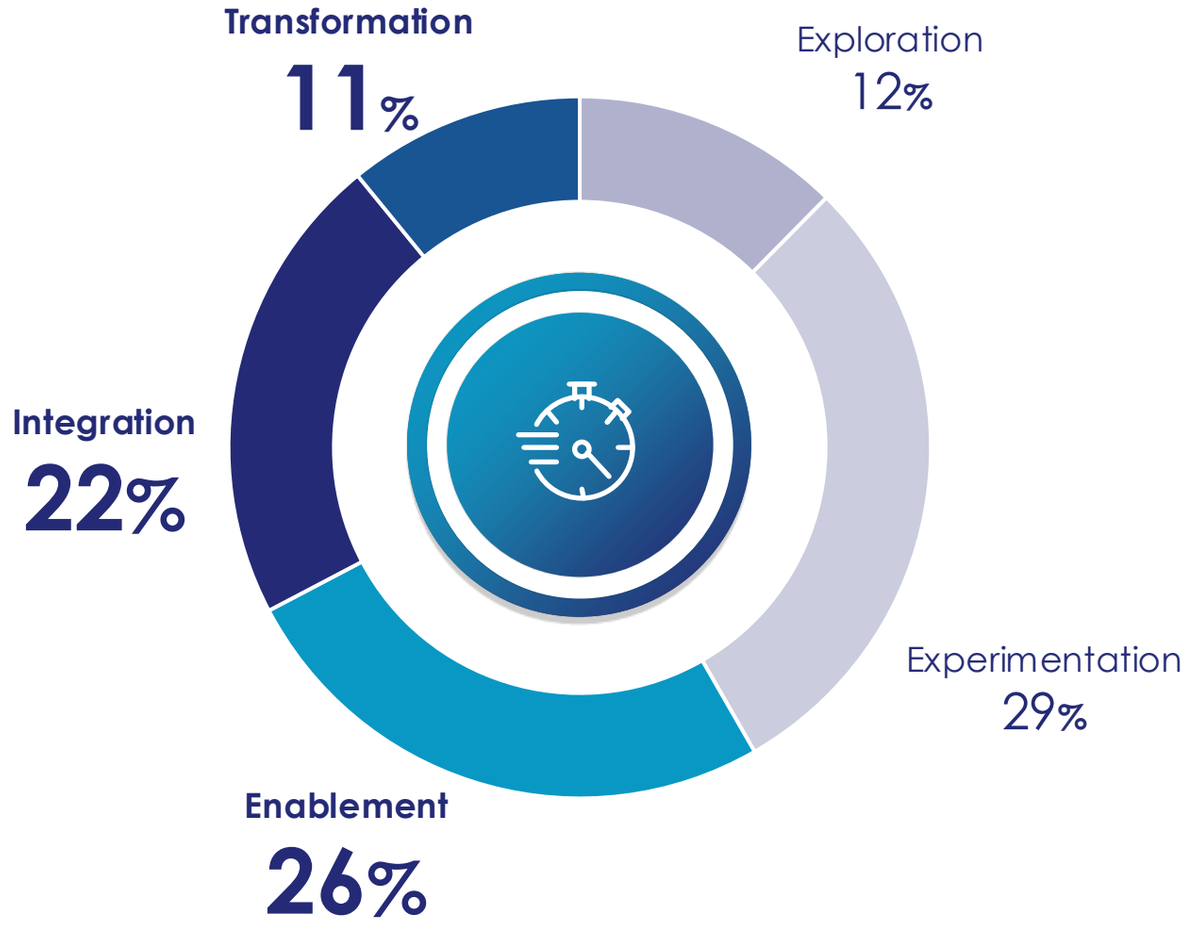
www.thalesgroup.com

AI adoption by enterprises speeds up

Wide adoption: AI is being adopted by enterprises around the world at a faster rate than other technologies such as internet or PCS.

Advanced deployment: A majority (59%) of organizations have moved from the early stages of experimentation and exploration to deployment stages of enablement, integration and transformation.

Transformation: 11% of organizations are leveraging AI to actively transforming their core business.[1]



1. *Thales Data Threat Report 2025*

Deployment of AI creates a new set of challenges to security teams

Ecosystem changes with new infrastructures, SaaS services and agents to support AI could represent new points of vulnerability.

Data integrity and trust challenges are exacerbated by AI and integrity attacks such as data poisoning, can inflict bias to AI models with incorrect data.

Confidentiality and compliance with regulations challenges increase with AI, where models and applications can inadvertently expose sensitive data and lead non-compliance.[1]

1. Thales Data Threat Report 2025

Attacks on AI models and applications leading to Data Loss are on the rise

https://incidentdatabase.ai (5000+ Detailed AI incident Reports)

Incident 1186: Reported Public Exposure of Over 100,000 LLM Conversations via Share Links Indexed by Search Engines and Archived

Description: Across 2024 and 2025, the share features in multiple LLM platforms, including ChatGPT, Claude, Copilot, Qwen, Mistral, and Grok, allegedly exposed user conversations marked "discoverable" to search engines and archiving services. Over 100,000 chats were reportedly indexed and later scraped, purportedly revealing API keys, access tokens, personal identifiers, and sensitive business data.

Two AI incident examples.
Lots of bad stuff going on here

- Training data containing sensitive information. (API keys, authentication credentials, user-names, email addresses, IP)

- LLMs storing user conversation history in publicly accessible logs.

- Search BOTs scraping LLM meta data (conversation history).

- API keys and access tokens leaked.

- Sensitive business data leaked. PII data, Intellectual Property & BI.

Incident 956: Alleged Inclusion of 12,000 Live API Keys in LLM Training Data Reportedly Poses Security Risks

Description: A dataset used to train large language models allegedly contained 12,000 live API keys and authentication credentials. Some of these were reportedly still active and allowed unauthorized access. Truffle Security found these secrets in a December 2024 Common Crawl archive, which spans 250 billion web pages. The affected credentials could have been exploited for unauthorized data access, service disruptions, financial fraud, and a variety of other malicious uses.

Top 10 OWASP Risks for LLMs and Gen AI Apps 2025

Attackers see AI as a new opportunity to gain access to sensitive information and are targeting AI initiatives across the entire lifecycle to find vulnerabilities that can be exploited.

LLM01 2025
Prompt Injection

A Prompt Injection Vulnerability occurs when user prompts alter the LLM's behavior or output in unintended ways.

LLM02 2025
Sensitive Information Disclosure

LLMs, especially when embedded in applications, risk exposing sensitive data, proprietary algorithms, or confidential details through their output.

LLM03 2025
Supply Chain

LLM supply chains are susceptible to various vulnerabilities, which can affect the integrity of training data, models, and deployment platforms.

LLM04 2025
Data and Model Poisoning

Data poisoning occurs when pre-training, fine-tuning, or embedding data is manipulated to introduce vulnerabilities, backdoors, or biases.

LLM05 2025
Improper Output Handling

Improper Output Handling refers to insufficient validation, sanitization, and handling of the outputs generated by LLMs before they are passed to other systems.

LLM06 2025
Excessive Agency

Excessive Agency enables damaging actions to be performed in response to unexpected, ambiguous or manipulated outputs from an LLM.

LLM07 2025
System Prompt Leakage

System prompt leakage refers to the risk that the LLM prompts or instructions used to steer the model's behavior can contain sensitive information that was not intended to be discovered.

LLM08 2025
Vector & Embedding Weaknesses

Vectors and embeddings weaknesses in how they are generated, stored, or retrieved can be exploited to inject harmful content, manipulate model outputs, or access sensitive information.

LLM09 2025
Misinformation

Misinformation occurs when LLMs produce false or misleading information that appears credible, leading to possible security breaches, reputational damage, and legal liability.

LLM10 2025
Unbounded Consumption

Unbounded Consumption refers to the process where a Large Language Model (LLM) generates outputs based on input queries or prompts.

https://genai.owasp.org/llm-top-10/

THALES
Building a future we can all trust

inception
Un evento de TrustDimension
EXECUTIVE SUMMIT
2026

AI Security Fabric

www.thalesgroup.com

AI risks stall AI adoption. The promise of AI value remains unrealized

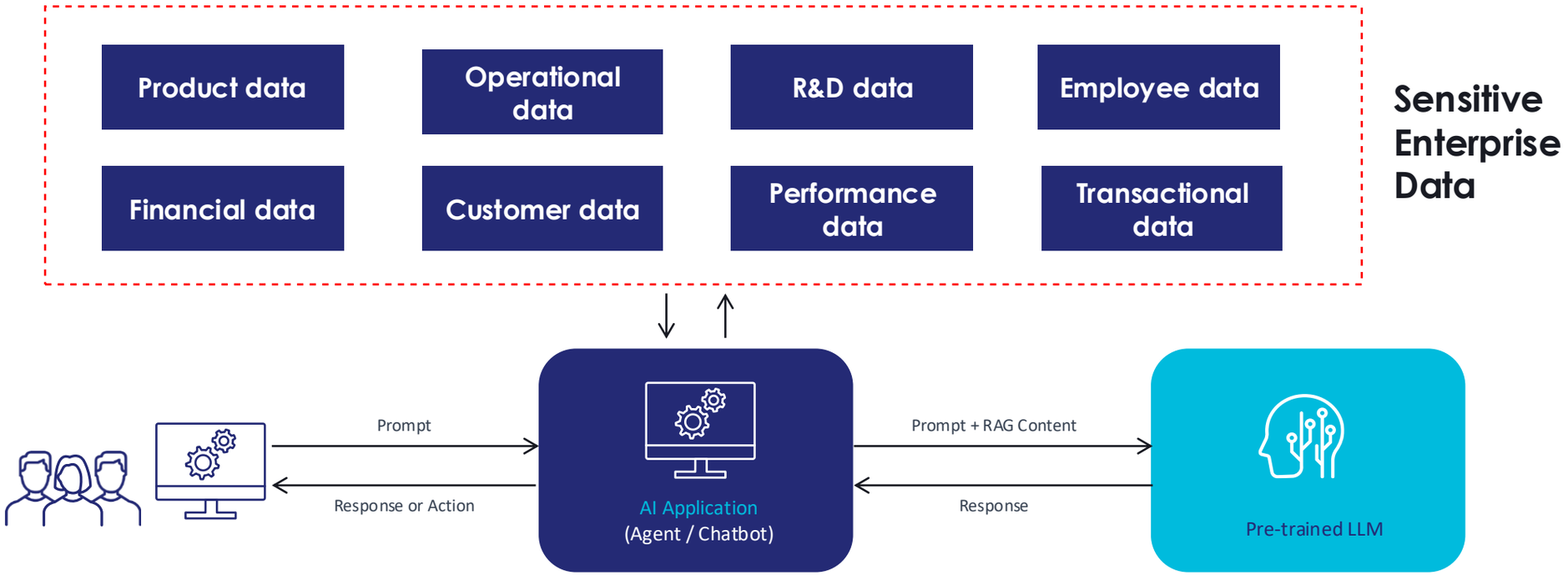
- AI leaders cite risk and compliance as top barriers to adoption, especially for agentic and sovereign AI, alongside regulatory monitoring and data residency constraints.
- Deloitte. AI trends 2025: Adoption barriers and updated predictions

- A McKinsey 2025 global survey finds nearly two-thirds of organizations have not yet begun scaling AI across the enterprise, despite experimentation.
- McKinsey. The State of AI: Global Survey 20

- Only 26% of organizations say they are realizing AI's full value, with the rest stuck in "pilot purgatory", according to BCG's executive surveys.
- BCG. AI at Work 2025: Momentum Builds, but Gaps Remain.

AI adoption implies access to your enterprise data

For AI to be truly effective for business, it must incorporate enterprise data that may include **sensitive and proprietary information.** The critical challenge is the **inherent security risk** of exposing this vital information.



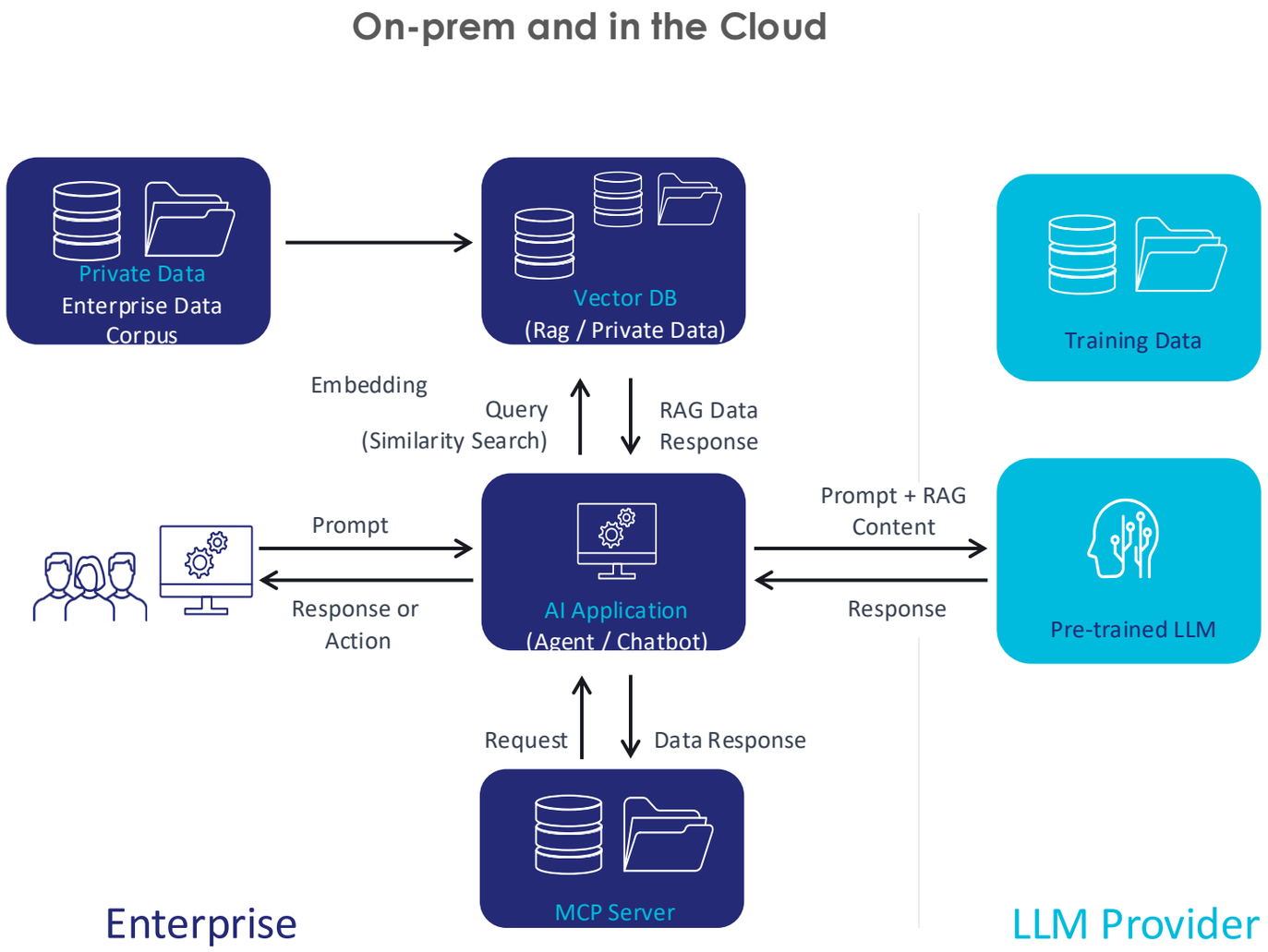
Examples include: Exposure to attacks, data leakage, unauthorized access, privacy concerns and regulatory failures.

AI generates new cyber risks

1. Prompt Data Leakage
2. Sensitive Data Training
3. Shadow AI
4. Data Re-Identification
5. Exposed AI APIs

> Data should be protected before being used by the training model to mitigate risk

AI Applications & Large Language Models (LLMs) simplified architecture



AI Security Fabric runtime security solutions for enterprises

Unlock the power of AI to drive business value while protecting data, applications and operations.

- Allow Agentic AI and Gen AI access to datasets while identifying and protecting sensitive data based on policy.
- Automate key security functions by enforcing custom policies based on security requirements and regulations such as GDPR, HIPAA or PCI.
- Address a majority of the OWASP Top 10 threats to LLMs by uncovering vulnerabilities, protecting data and applications, and detecting and stopping malicious activity in real time.
- Proven enterprise-ready solutions supported by decades of security expertise and based on globally deployed cybersecurity products.

AI Security Fabric runtime security solutions

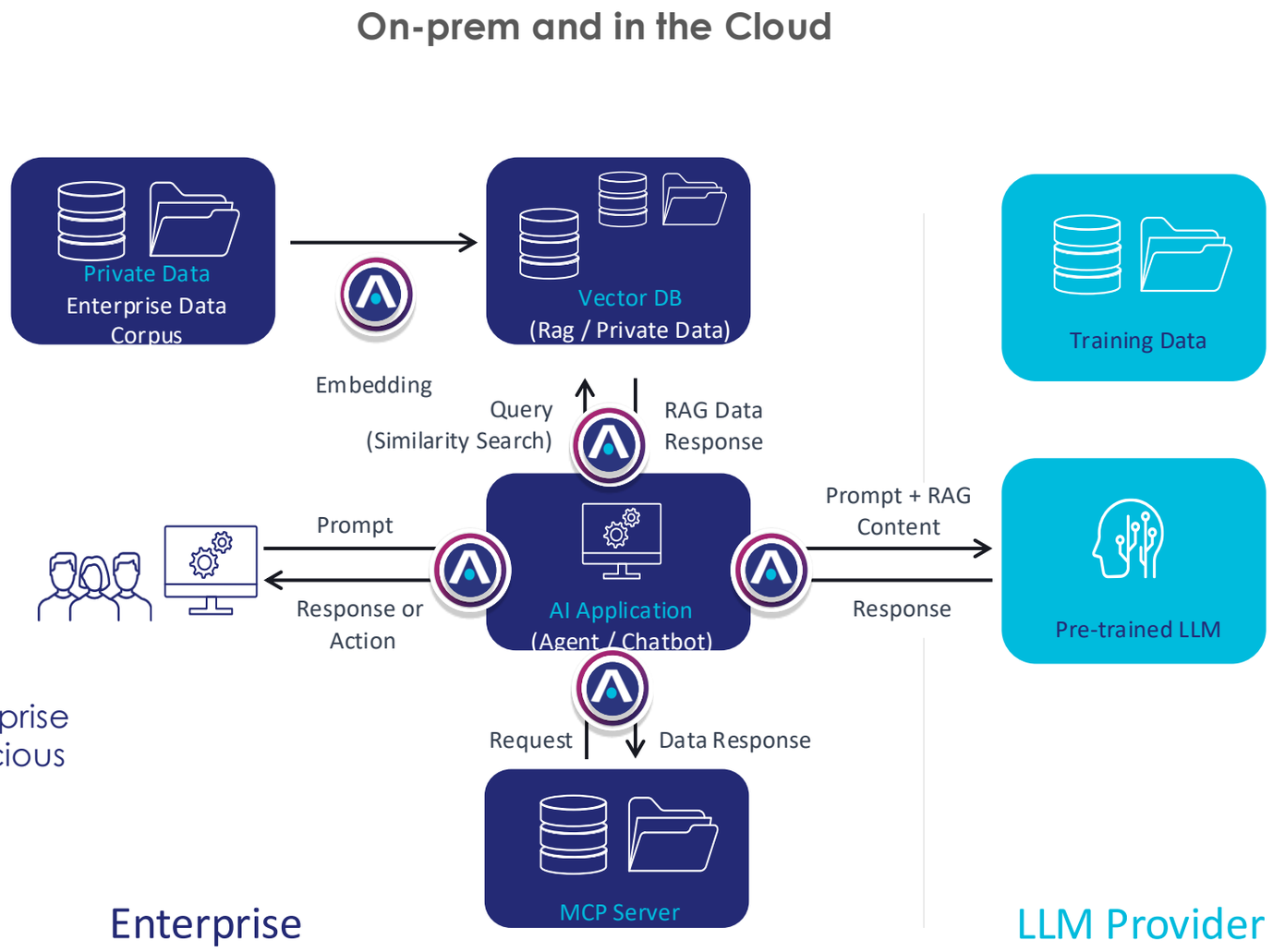
AI Application Security

Enterprise-grade protection for GenAI applications

RAG Data Protection

Safeguard sensitive data in an AI RAG systems

AI Runtime Security Solutions within overall AI architecture

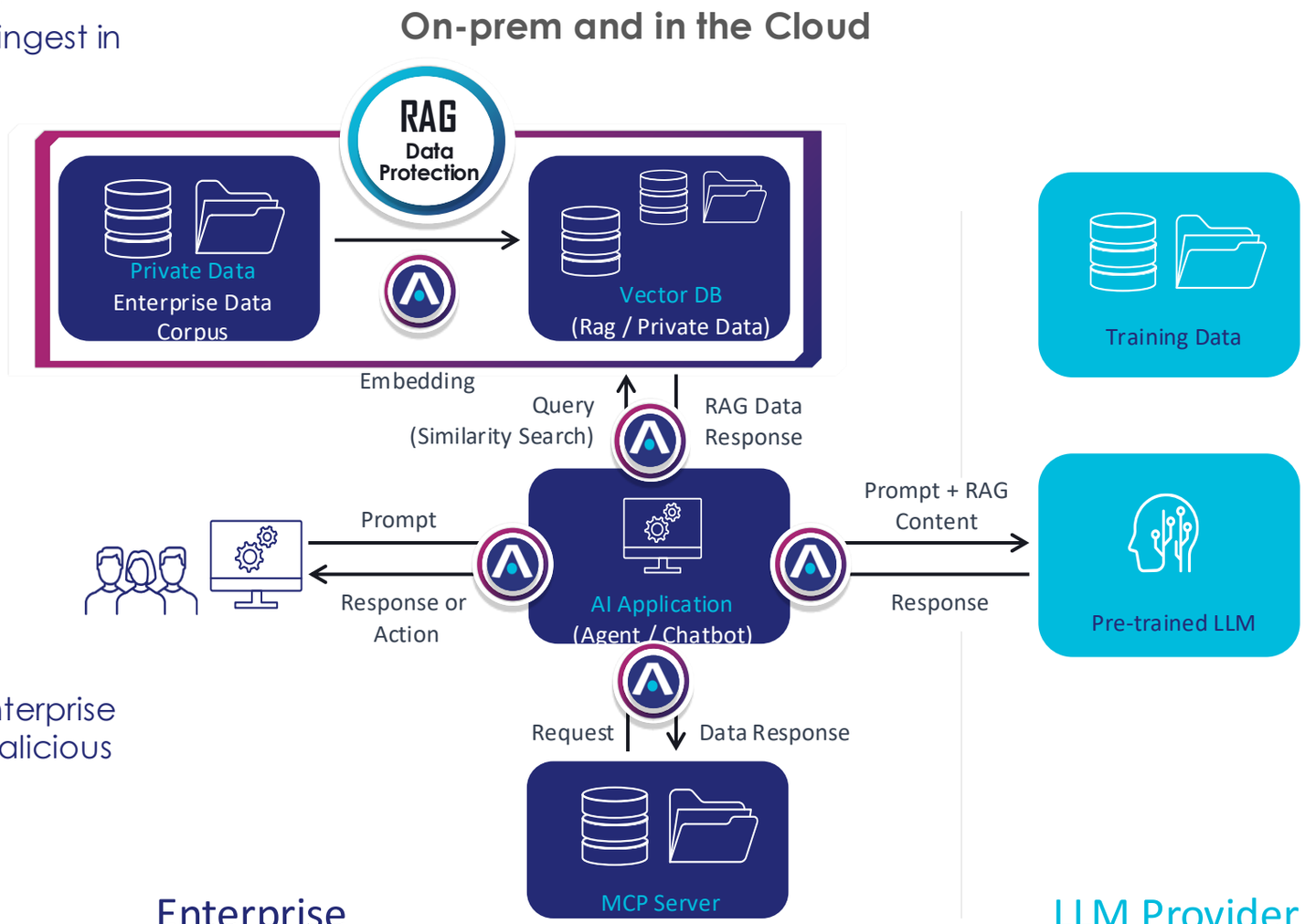


AI Firewall - Protect Enterprise AI Application from Malicious Prompts

AI Runtime Security Solutions within overall AI architecture

RAG Data Protection - Safeguard your data ingest in RAG system

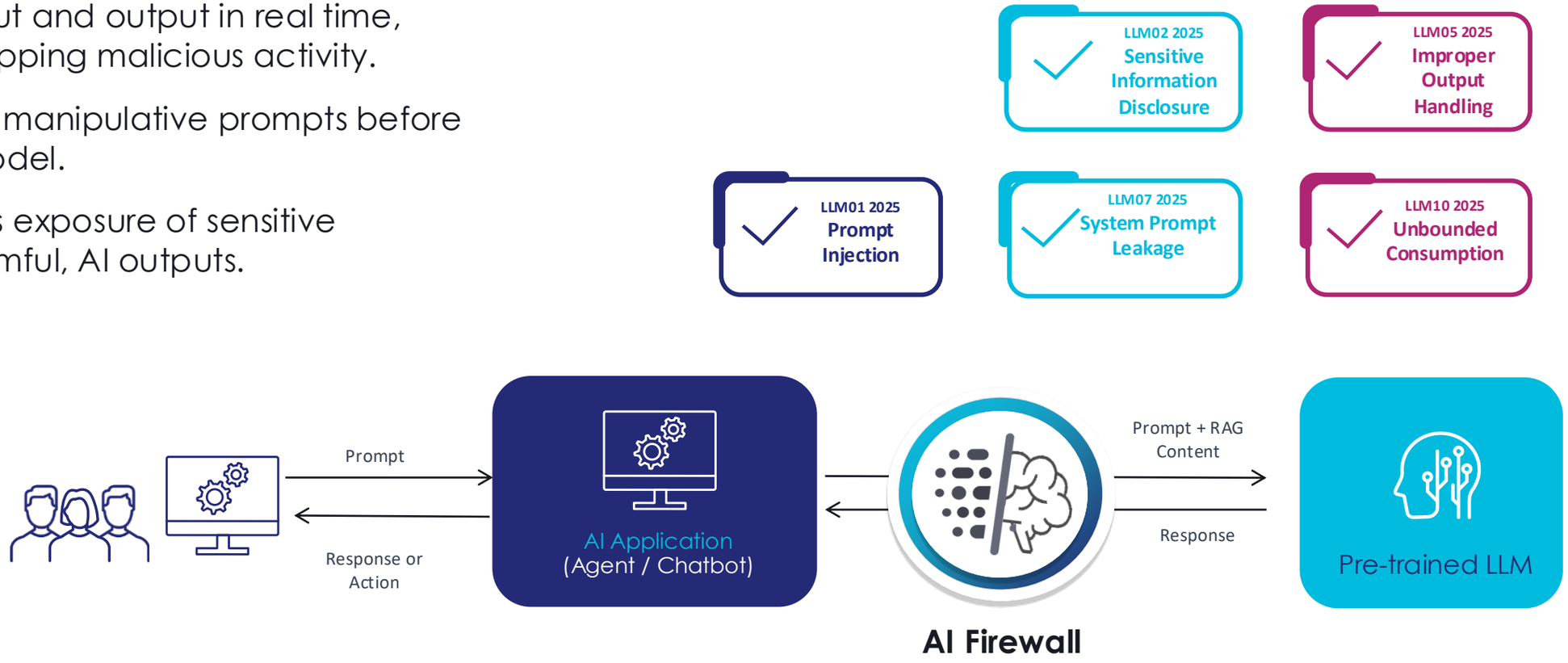
AI Firewall - Protect Enterprise AI Application from Malicious Prompts



Enterprise-grade protection for GenAI applications

The Imperva AI Firewall is a SaaS reverse proxy that sits between your applications and LLMs they rely on, analyzing all inputs sent to the models and all outputs they produce - while preserving application performance.

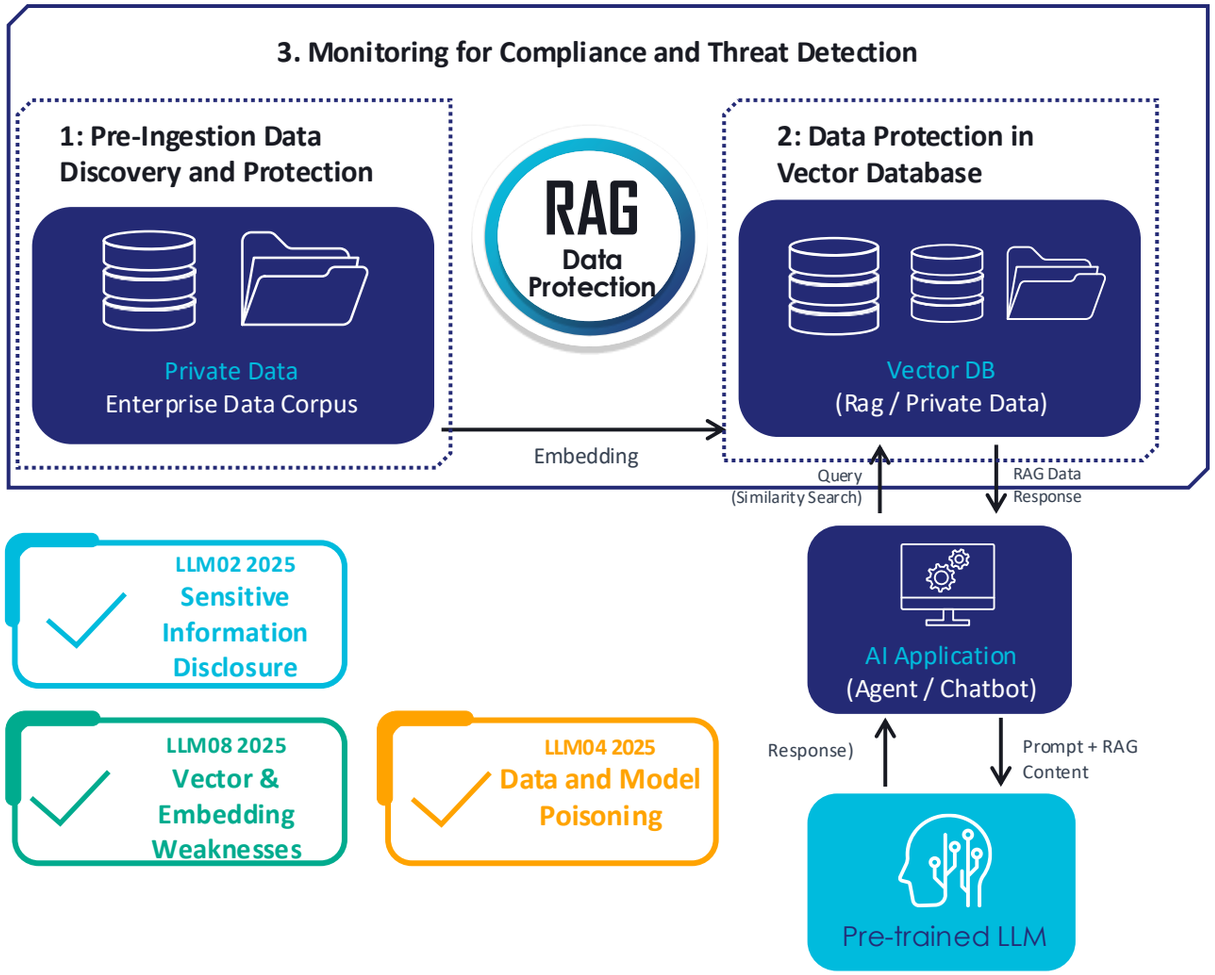
- Analyze every input and output in real time, detecting and stopping malicious activity.
- Block malicious or manipulative prompts before they reach the model.
- Detect and blocks exposure of sensitive information or harmful, AI outputs.



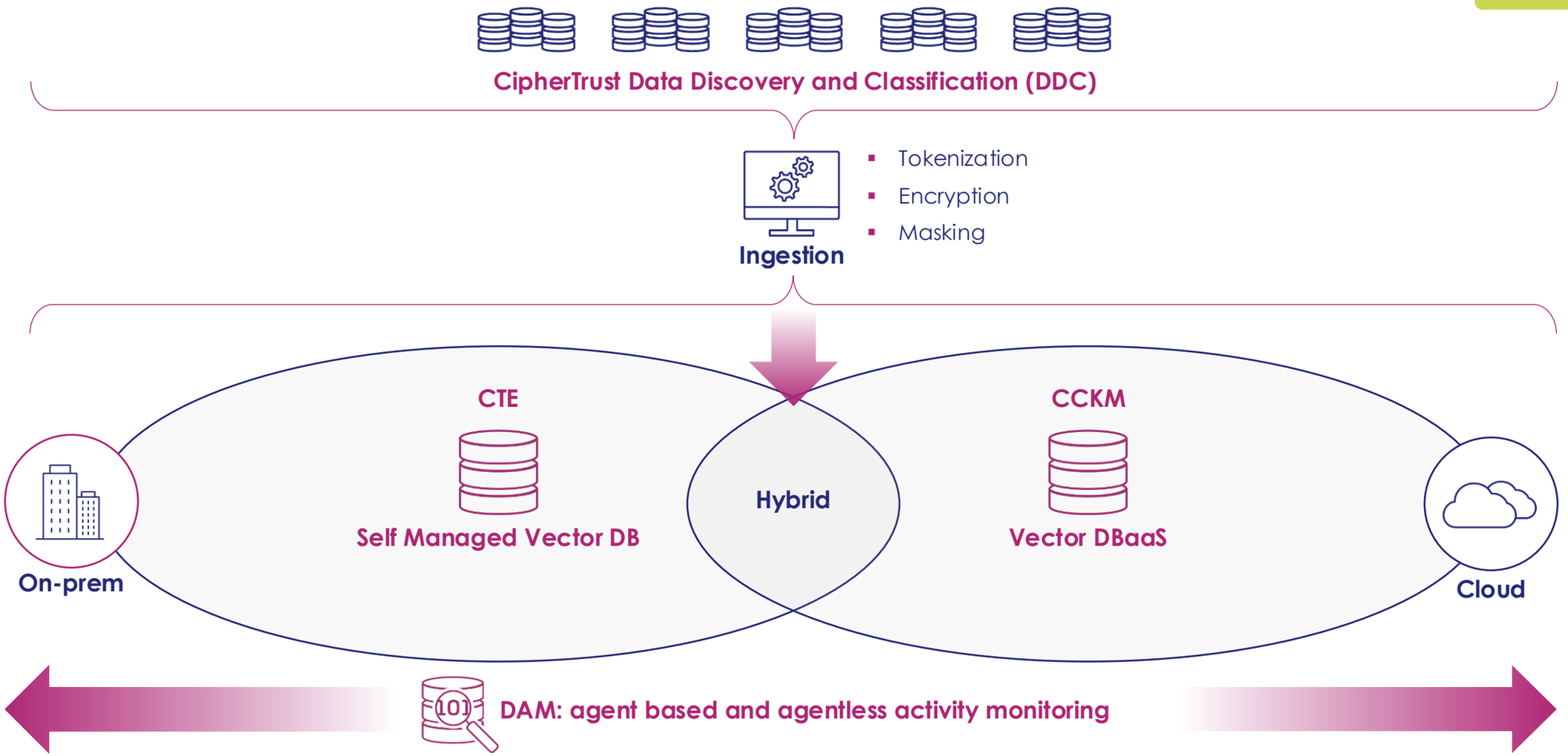
Safeguard sensitive data in an AI Retrieval-Augmented Generation (RAG)

Thales provides a comprehensive suite of **Retrieval-Augmented Generation (RAG) Data Protection Solutions** to safeguard sensitive data throughout its lifecycle within an enterprise AI application's RAG systems.

1. Identify, classify and protect sensitive data stores prior to ingestion into AI RAG systems.
2. Protect data during the ingestion, embedding and within the vector database.
3. Continuous, real-time monitoring of all interactions with databases and unstructured data across the RAG lifecycle.



Thales RAG Coverage



THALES
Building a future we can all trust

inception
Un evento de TrustDimension
EXECUTIVE SUMMIT
2026

Data Security Posture Management for AI

www.thalesgroup.com

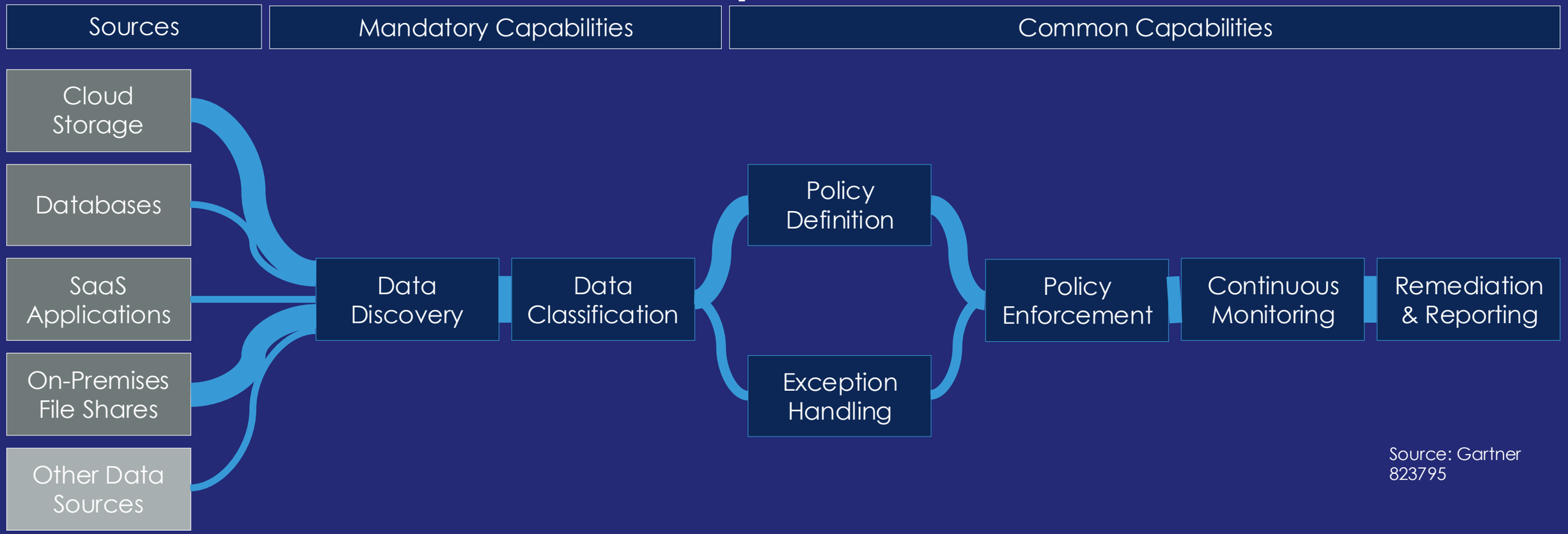
What is DSPM: High-level

Data Security Posture Management (DSPM) is a process and framework that helps organizations identify, assess, and manage data security risks across multiple environments. It protects sensitive data from unauthorized access, misuse, and/or theft**.** Often, DSPM is used by organizations to meet privacy and security regulations, prevent data breaches and exfiltration, and avoid other cyber attacks such as malware and ransomware. DSPM Focuses on the following cybersecurity practices to safeguard data:

Benefits of DSPM

1. **Easily identify** where data, especially sensitive data, resides across the entire data environment
2. **Quickly address data security and compliance risks**
3. **Identify precursors and vulnerabilities** that can lead to data breaches, exfiltration, and other cyberattacks to prevent impact on the organization and business.

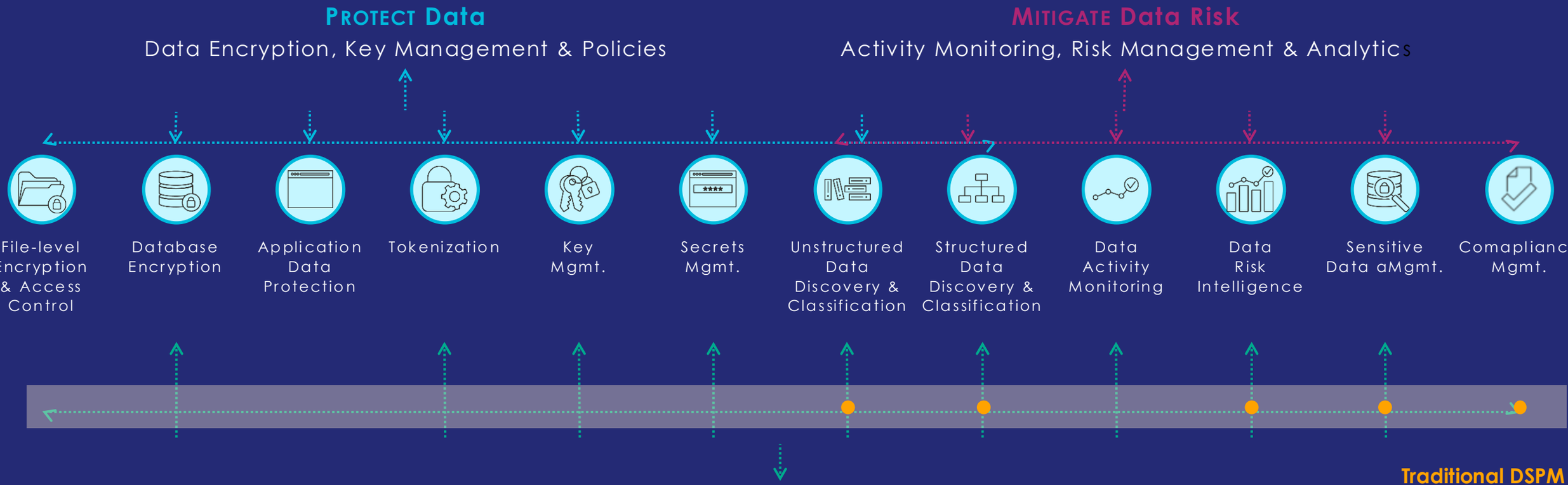
Gartner's Definition of DSPM Capabilities



Source: Gartner 823795

Thales Ciphertrust Data Security Platform + Data Security Fabric

Expanding our focus to Data Security Posture Management (DSPM)



"Thales offers superior capabilities for data discovery, Encryption, Tokenization, and data-access controls, including visibility of cryptographic posture and governance of keys and secrets."

Forrester Wave ™ Data Security Platforms Q1 2025 (Mar 2025)

DSPM - Focused on Minimizing Data Risk

Discover:

CONTINUOUSLY DISCOVER & CLASSIFY DATA AT SCALE

Discover & Classify Data:
Automatically discover all unstructured data, structured data and sensitive data stores(on-premises, hybrid, and multi-cloud environments)

Monitor Data Access & Activity:
Complete visibility with continuous monitoring, auditing, and analysis of all data stores and data types.

Analyze:

INTELLIGENCE TO ACCURATELY ASSESS DATA RISK

Analyze Risk:
Analyzes data sensitivity, access, usage, and security controls to quantify real data risk.

Prioritize Risk:
Combine key data risk indicators into an actionable view to understand where and why your data is at risk and address attack vectors quickly.

Protect:

INSTANT, CONFIGURABLE DATA DEFENSES

Deploy Data Protection:
Secure, anonymize, and encrypt data at rest and in motion across the entire IT ecosystem.

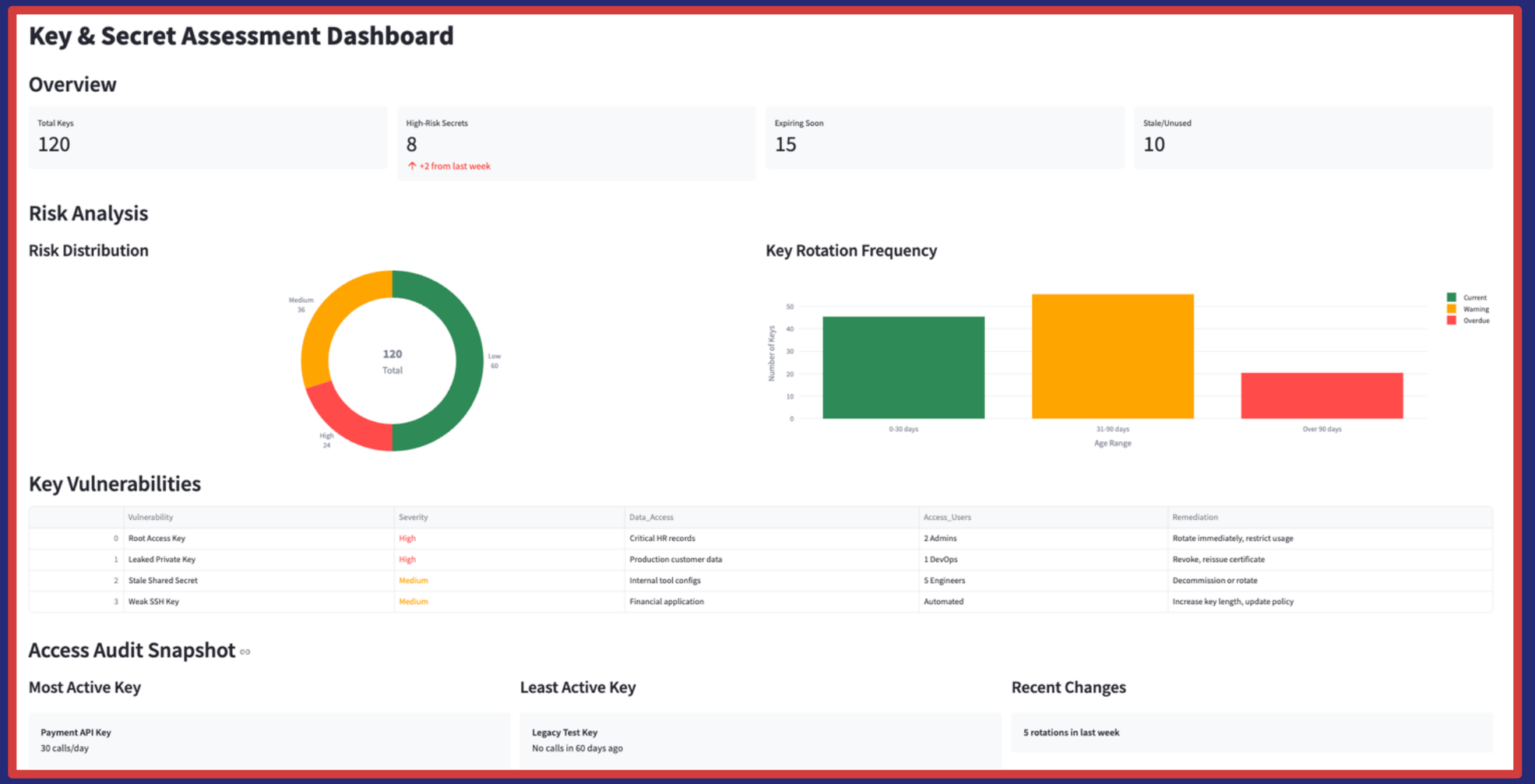
Control:

CENTRALLY MANAGED PRIVILEGES AND DATA ACCESS

Manage Data Credentials:
Manage cryptographic keys, policies, and secrets centrally to prevent unauthorized access and lessen credential attacks.

Align to Compliance:
Automate and simplify regulatory compliance activities, providing superior long-term retention of live audit data.

Plus Keys and Secrets Security Posture Management

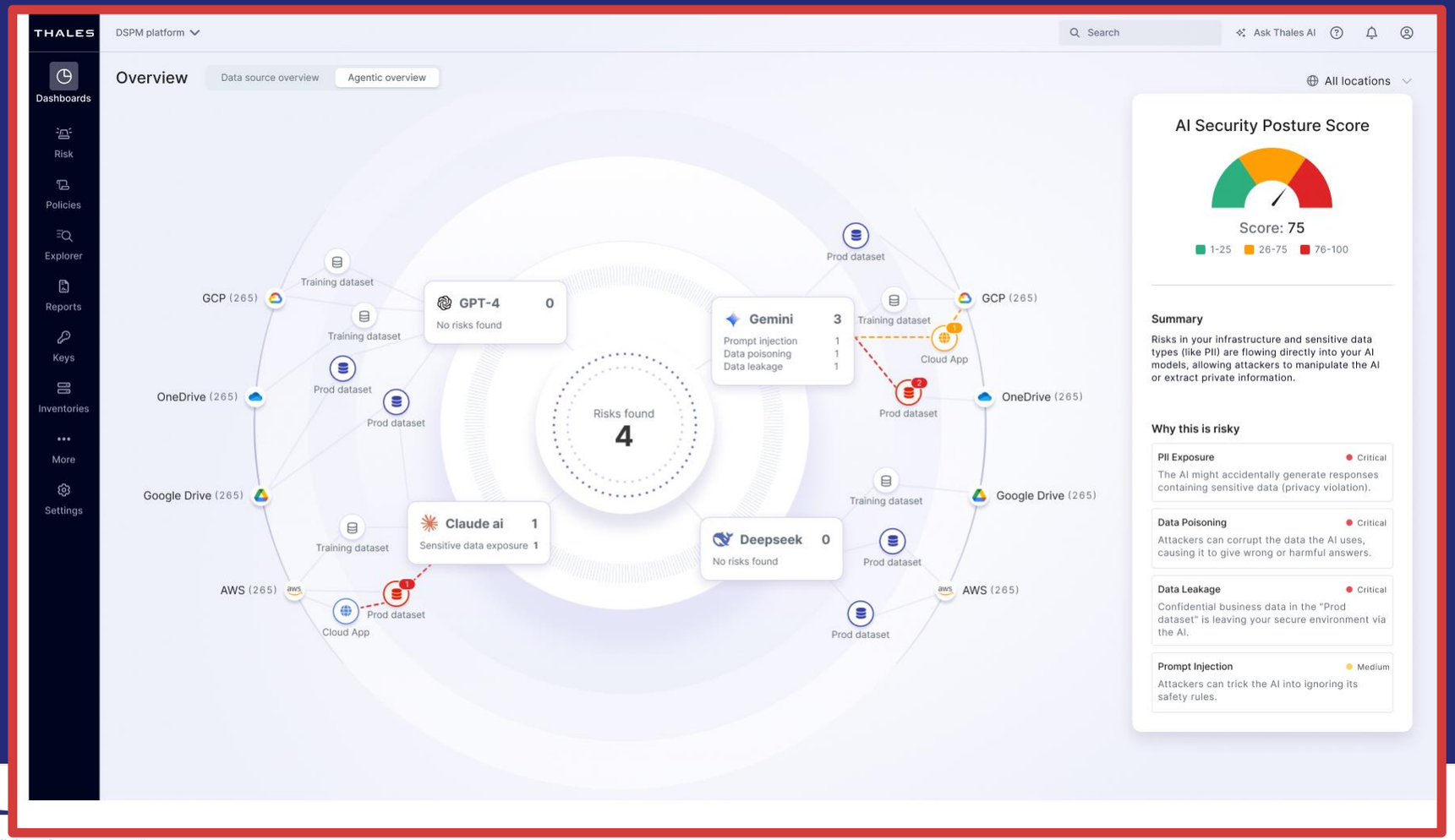


Centralized discovery & classification

Deeper risk assessment for keys & secrets

AI Posture Management

Extending our DSPM platform to become a comprehensive data security solution protecting ALL data in the organization -including data accessed and/or data stored in agentic workforce



- AI asset discovery & mapping of connections between assets
- Data classification
- Access and permission monitoring
- Risk analysis & mitigation

THALES
Building a future we can all trust

inception
Un evento de TrustDimension
EXECUTIVE SUMMIT
2026

Conclusions

www.thalesgroup.com

Thales solutions can help address 7 of the top 10 OWASP for LLMs and Gen AI Apps 2025

LLM01 2025
Prompt Injection

Block malicious prompts and injections before they reach your models

LLM02 2025
Sensitive Information Disclosure

Identify and protect sensitive data before reaching the model

LLM03 2025
Supply Chain

LLM supply chains are susceptible to various vulnerabilities, which can affect the integrity of training data, models, and deployment platforms.

LLM04 2025
Data and Model Poisoning

Reduce the risk of data poisoning and model manipulation by protecting training data

LLM05 2025
Improper Output Handling

Detect and filter harmful or inappropriate content in user interactions during conversations

LLM06 2025
Excessive Agency

Excessive Agency enables damaging actions to be performed in response to unexpected, ambiguous or manipulated outputs from an LLM.

LLM07 2025
System Prompt Leakage

Monitor output and guard against system prompt exposure and improper output

LLM08 2025
Vector & Embedding Weaknesses

Secure data stored in vector database with encryption and key management

LLM09 2025
Misinformation

Misinformation occurs when LLMs produce false or misleading information that appears credible, leading to possible security breaches, reputational damage, and legal liability.

LLM10 2025
Unbounded Consumption

Prevent malicious AI resource use and denial of service attacks intended to drive up costs

https://genai.owasp.org/llm-top-10/

Thales AI Runtime Security: Enabling innovation, securing data

Accelerate AI adoption and results by securely leveraging your enterprise data

Unlock business value by deploying AI with confidence

Harness the power of AI without creating vulnerabilities for users, operations, or data.

Improve AI Security Posture Management

Improve AI Data Security Posture Management by gaining visibility on data, apps, users and their interactions with AI Models and applications.

Prevent costly and incidents and maintain compliance

Identify and block common prompt injection techniques and reduce risk of data leakage, and model manipulation before they impact your business.

Tools available for you:

- AI Security Workshop
- AI Firewall demo / test
- DSPM Assessment
- Cloud Keys and Secrets Discovery
- Imperva for Google Cloud integration

Thank you!

www.thalesgroup.com