RAG data protection risks and challenges

cpl.thalesgroup.com

THALES
Building a future we can all trust

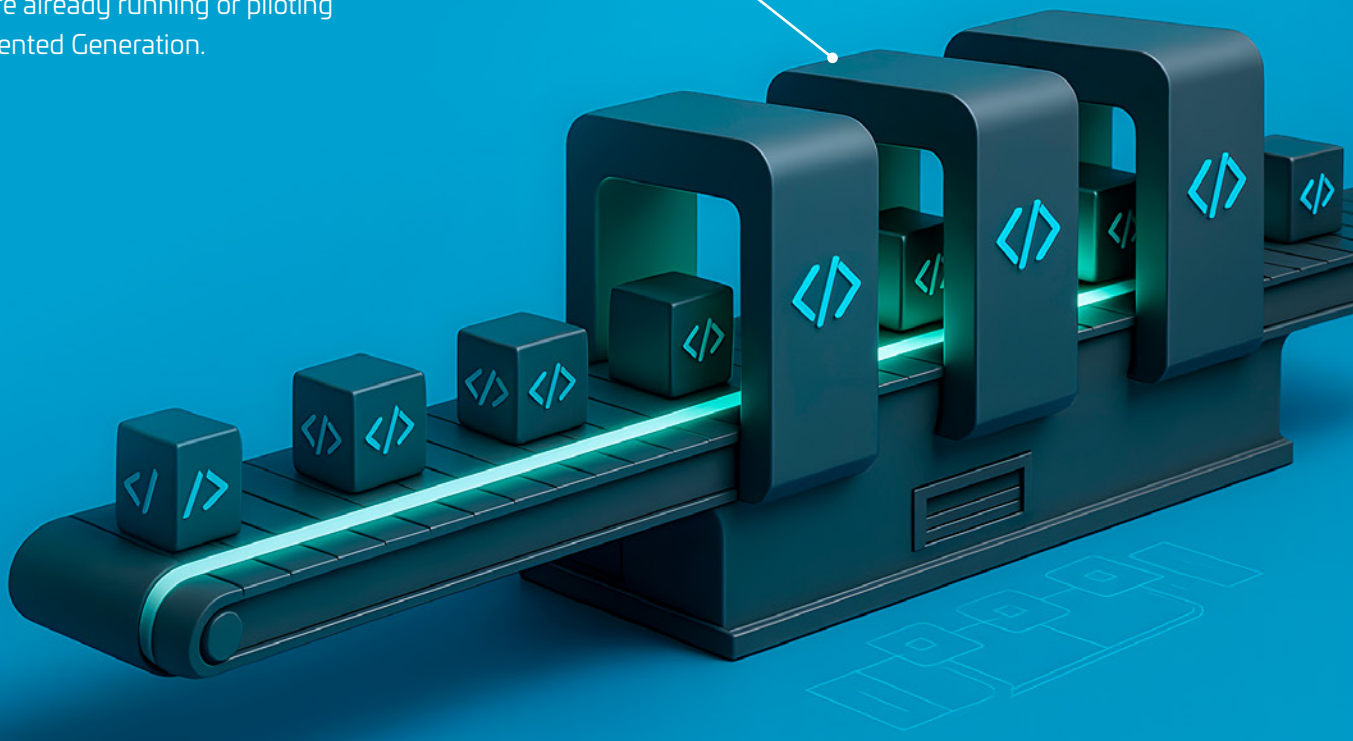
THALES

CYBERSECURITY

RAG Gives AI a Memory. Most Organizations Aren't Protecting It.

51%

of enterprises are already running or piloting Retrieval-Augmented Generation.



RAG is how enterprises make AI actually useful. Instead of relying on generic training data, RAG connects LLM models to your internal knowledge: contracts, patient records, financial data, source code, and customer files. That's powerful, but it's also a new attack surface that most security programs weren't built to handle.

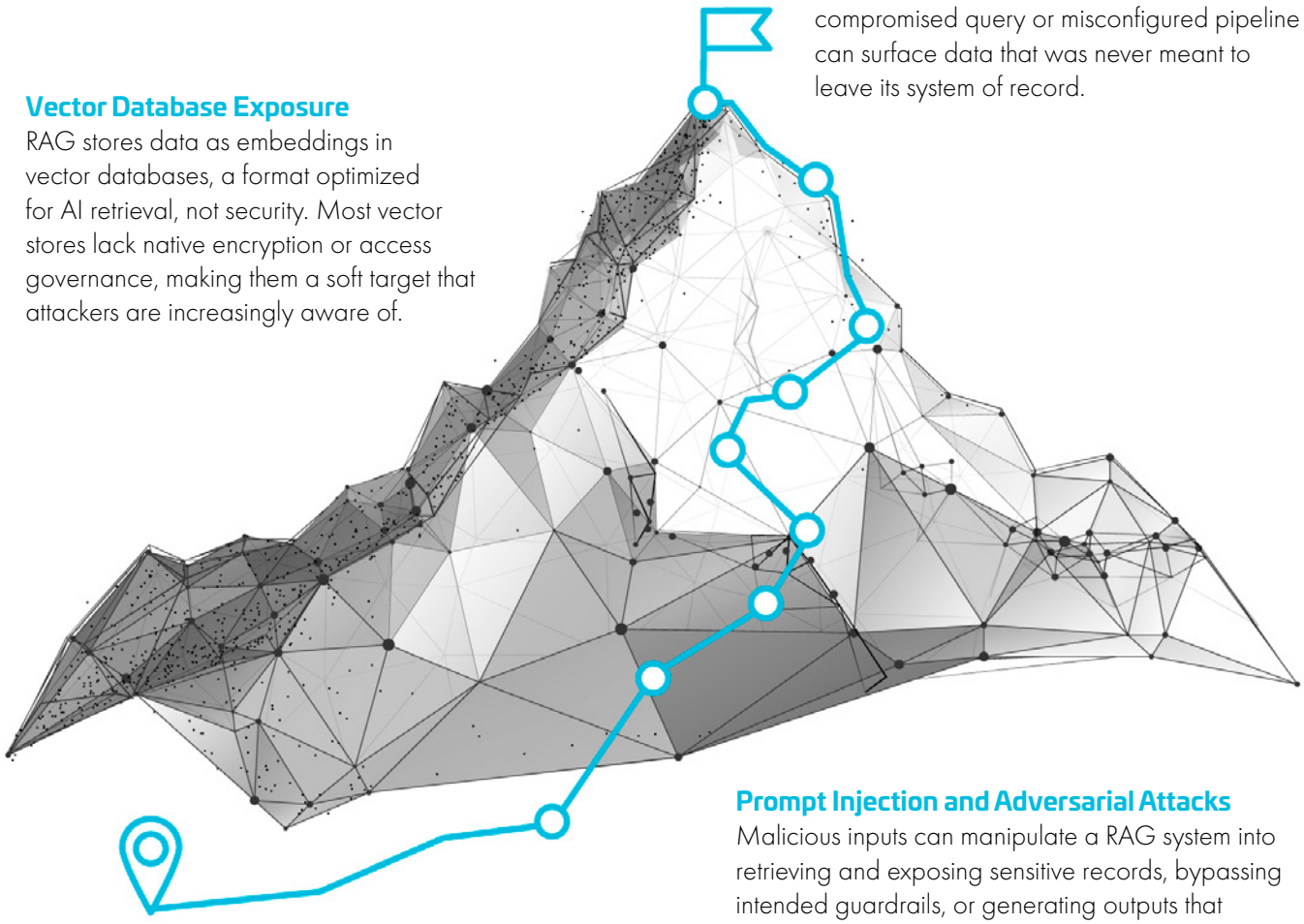
Security Risks

Unauthorized Data Access

RAG systems retrieve data on demand, often without the fine-grained access controls applied to human users. A single compromised query or misconfigured pipeline can surface data that was never meant to leave its system of record.

Vector Database Exposure

RAG stores data as embeddings in vector databases, a format optimized for AI retrieval, not security. Most vector stores lack native encryption or access governance, making them a soft target that attackers are increasingly aware of.



Prompt Injection and Adversarial Attacks

Malicious inputs can manipulate a RAG system into retrieving and exposing sensitive records, bypassing intended guardrails, or generating outputs that misrepresent your data. 61% of organizations report their AI applications are being actively targeted by attackers. These attacks are difficult to detect and are getting more sophisticated.

Data Leakage Through AI Outputs

When sensitive data is retrieved and included in a generated response, it can travel far beyond its original authorization boundary: shared in chat interfaces, exported in reports, or surfaced to users who lack clearance to the source data.

OPERATIONAL & ADMINISTRATIVE CHALLENGES

You can't protect what you can't see.

Only 34% of organizations have complete knowledge of where their data is stored

RAG ingests across structured and unstructured sources: databases, file shares, cloud storage, and collaboration tools. Without pre-ingestion discovery and classification, sensitive data flows into AI pipelines before anyone has assessed the risk.

Unstructured data is the hardest to govern.

80% of enterprise content is unstructured

This is exactly the data RAG draws from. It's also the data least likely to have consistent labeling, access policies, or retention controls applied.

Encryption key ownership is ambiguous.

A majority of organizations allow cloud providers to control encryption keys for more than half of their applications.

In a RAG context, that means your AI's data store may not be under your cryptographic control, creating gaps in both your security posture and regulatory compliance.

No audit trail for AI data interactions.

34% of organizations have experienced AI-related data breaches

RAG creates a new category of data interaction: automated, high-velocity, and often invisible to existing monitoring tools. Without purpose-built logging, organizations can't answer basic questions: what data did the AI access, when, and why?

Compliance frameworks weren't written for RAG.

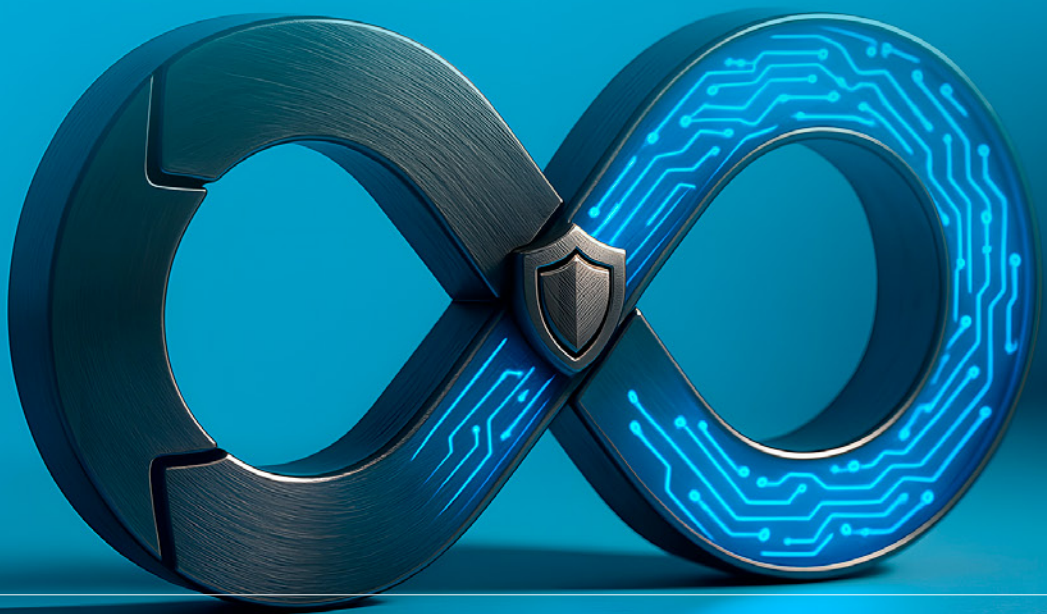
Organizations that failed a compliance audit had only a 6% no-breach rate, compared to 30% among those that passed all audits

GDPR, HIPAA, PCI DSS, and sector-specific regulations govern how data is stored, accessed, and shared. RAG systems introduce retrieval patterns, output channels, and data residency questions that these frameworks didn't anticipate, and that compliance teams are still learning to address.

Tool fragmentation makes it worse.

Organizations average 7 data protection and monitoring tools.

None were designed with RAG in mind. Bolting existing controls onto a RAG pipeline creates gaps and misconfigurations leading to costly breaches.



THE NUMBERS BEHIND THE RISK:

6 data points that show how fast the gap is growing.

- **51%** of enterprises are already running or piloting RAG
- **80%** of enterprise content is unstructured — the primary fuel for RAG systems
- **34%** of organizations have experienced AI-related data breaches
- **Only 34%** have complete knowledge of where their data is stored
- **61%** report their AI applications are being actively targeted by attackers
- Organizations average 7 data protection and monitoring tools

RAG doesn't just connect your AI to your data.
It connects every risk in your data estate to your AI.

Want to see how RAG data protection works in practice?
Watch the overview

THALES

CYBERSECURITY

Contact us

For contact information, please visit https://cpl.thalesgroup.com/contact-us

cpl.thalesgroup.com