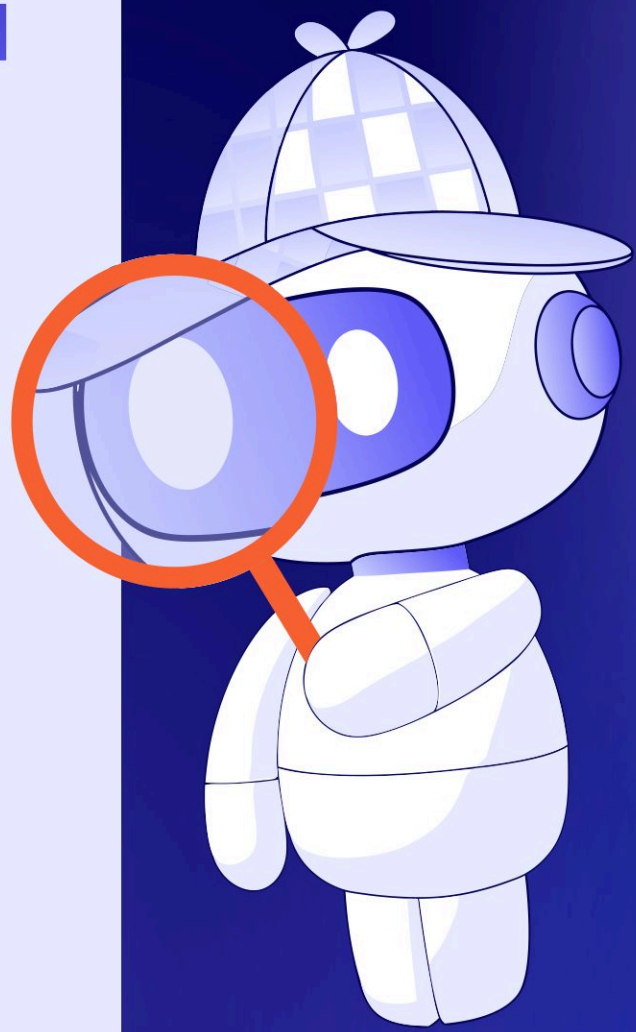
CASE STUDY

Where the Data Was Hiding



Real data. Real risks. Real results.

Sensitive data has a way of ending up in places nobody intended, and AI is making that much harder to ignore.

Every tool an employee picks up can read whatever that employee can reach, and most companies would struggle to say what sits in that reach or who else already has access to it.

Concentric AI helps organizations find sensitive data hiding in plain sight, understand who can access it and how it's used, and take action to reduce risk.

Here are just a few examples of what we've uncovered.

Investment and Capital Markets

The challenge: A boutique investment firm with 40 years of history was deploying Office 365 and OneDrive while a small security team managed sensitive content using folder-level permissions and rules.

What Concentric AI uncovered: Six million files scanned in a single day, organized into 200 thematic clusters, with roughly 300 documents flagged for containing business-sensitive content and permissions that warranted attention.

The result: A clear view of sensitive data and access risk that would have been difficult to identify using folder-level controls alone.

Banking and Credit Unions

The challenge: A Texas credit union serving hundreds of thousands of members and managing billions of dollars in assets had failed three separate DSPM rollouts because the vendors relied on regular expressions to identify sensitive data.

What Concentric AI uncovered: Member data across on-premises file shares and OneDrive, identified using content and context rather than regex-based detection.

The result: The security team had usable classification results by the second day of deployment.

Asset Management

The challenge: A 400-person investment management firm had blocked AI company-wide after a peer firm leaked confidential data into a public chatbot.

What Concentric AI uncovered: Visibility into prompt activity across Copilot, ChatGPT, and Perplexity, along with sensitive-data controls and approval workflows to identify risky uploads before they leave the organization.

The result: A practical approach to governing AI use without relying on a blanket AI ban.

Education

The challenge: A university with 25,000 students had moved transcripts, academic records, and financial aid data to SharePoint, OneDrive, and Azure during its transition to hybrid learning—with no clear visibility into who had access.

What Concentric AI uncovered: Every repository was scanned within a day, including unused accounts that still had access to FERPA-regulated records.

The result: The university identified access risks and estimated $2 million in annual savings by avoiding potential penalties and administrative costs.

Healthcare

The challenge:
A Fortune 100 medical device manufacturer had PHI and clinical trial data scattered across databases and file-sharing platforms used to share with hospitals, research partners, and third-party vendors.

What Concentric AI uncovered:
Sensitive data across the environment without rule writing or regex tuning.

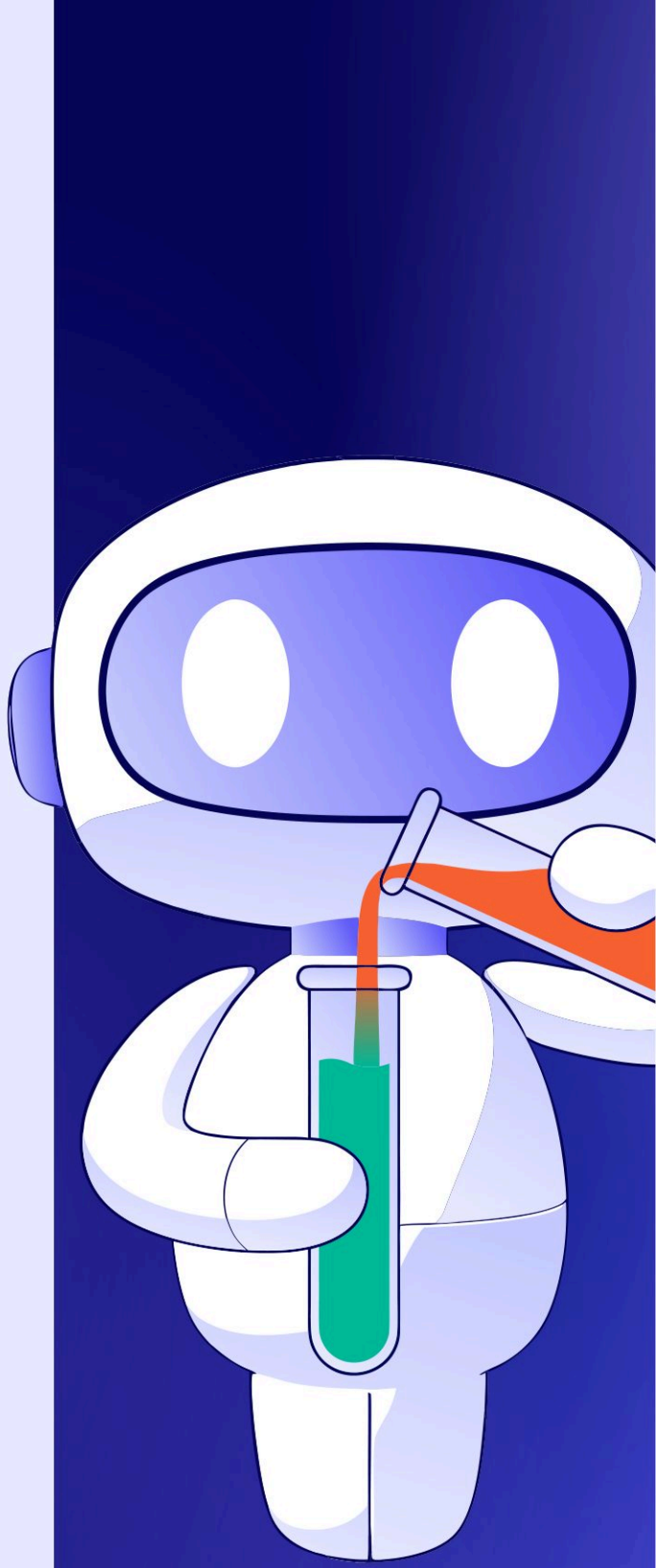
The result:
Categorization and risk assessment time was reduced by 60%, with an estimated 75% improvement in accuracy of HIPAA and GDPR compliance.

Public Health

The challenge:
A regional public health agency processing thousands of COVID tests each day was running six servers to support a competitor's classification tool, which still couldn't detect inappropriate access to PHI.

What Concentric AI uncovered:
Unprotected health data on shared drives during the proof of concept.

The result:
Concentric AI replaced the six-server deployment with a single server, reducing time spent on compliance and governance by more than 75%.



Energy

The challenge: A Fortune 1000 oil, gas, and aerospace services company relied on manually applied Controlled Unclassified Information (CUI) labels across siloed departments, creating ITAR compliance challenges for NASA and defense contracts.

What Concentric AI uncovered: CUI across the organization's environments without relying on manual classification.

The result: Automated CUI discovery and classification, along with controls that blocked SharePoint sharing to non-business email addresses—shifting the organization from reactive cleanup to proactive governance.

Technology

The challenge: A 10,000-person electronic systems design company had already experienced trade-secret loss involving departing employees. Reviewing SIEM and endpoint logs for insider activity required the equivalent of 10 person-months.

What Concentric AI uncovered: A senior engineer uploading proprietary algorithms to cloud storage during their final week at the company.

The result: Insider risk reviews that previously took 30–60 days were reduced to hours.

Manufacturing

The challenge: A large food manufacturer relied on a legacy data security tool that had been broken for three months, leaving generations of proprietary process knowledge unmonitored.

What Concentric AI uncovered: Risky user behavior involving sensitive proprietary data, including a departing employee sending a large volume of company data to themselves.

The result: The organization replaced its legacy tool with a data categorization and user behavior analytics solution that surfaced activity requiring investigation.

Retail

The challenge:
A mid-market outdoor apparel and workwear retailer rolling out Microsoft Copilot lacked a clear view of sensitive data across Google Drive, OneDrive, SharePoint, and file shares—or of who could access it via external links and stale permissions.

What Concentric AI uncovered:
20 TB of unstructured data classified according to the company's own policy categories, along with permission risks, including links open to anyone.

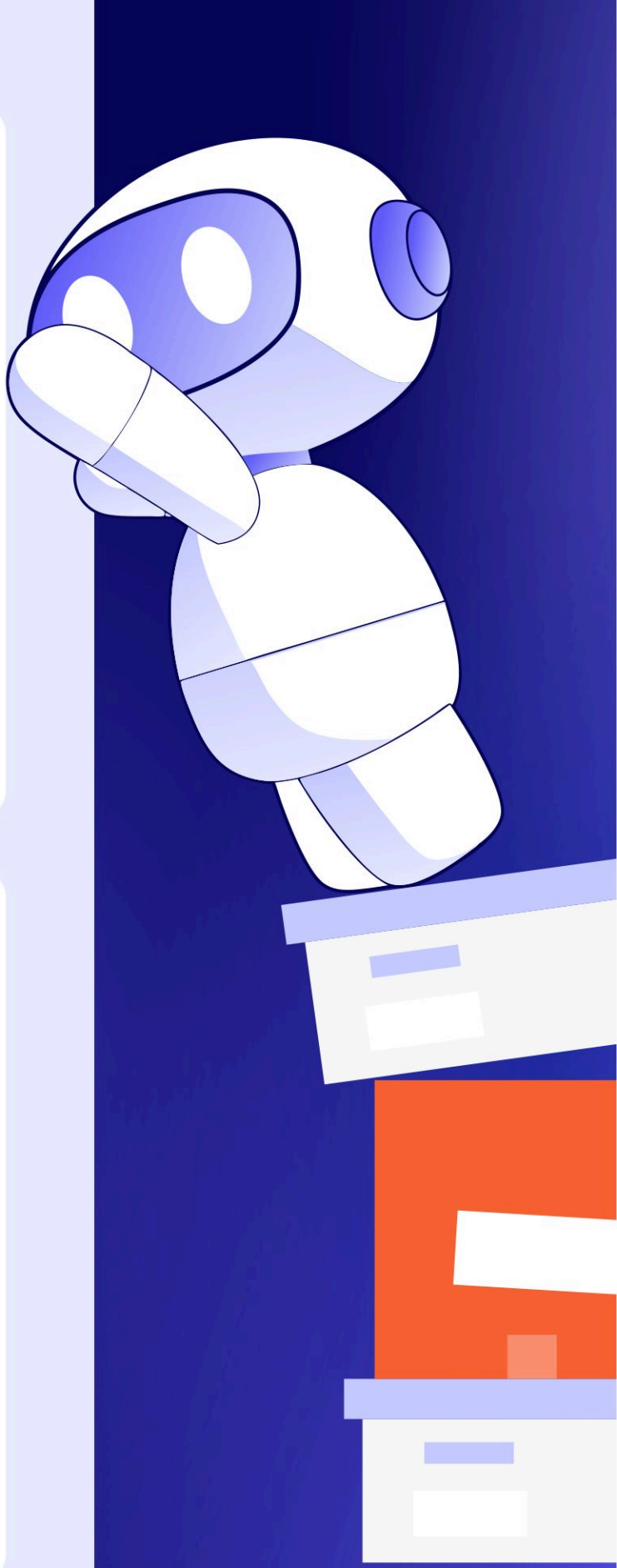
The result:
A stronger data foundation for Copilot, paired with Gen AI DLP to govern employee use of unsanctioned AI tools.

Professional Services

The challenge:
A 200-person actuarial and pension administration firm had accumulated decades of poor-quality records through acquisitions, with much of it stored as scanned PDFs that were difficult to analyze at scale.

What Concentric AI uncovered:
Sensitive information in legacy documents and access patterns spanning cloud and on-premises storage.

The result:
The firm established its first workable data retention policy and secured stronger support for SOC 2 compliance.



Insurance

The challenge: A national benefits carrier and absence management administrator processes millions of disability and leave claims, many arriving as scanned handwritten forms in PDFs.

What Concentric AI uncovered: U.S. Social Security numbers handwritten in scanned documents—data that pattern matching, text extraction, and traditional OCR could not reliably identify.

The result: Visibility into sensitive information hidden within unstructured, hard-to-read documents.

The challenge: A national life insurance company within a global insurance group had 50 TB of unstructured files across file shares and SharePoint, with Social Security numbers, dates of birth, and salary data buried in documents that had gone untouched for five years or longer.

What Concentric AI uncovered: More than 20 million files classified, including thousands of SSN and passport records.

The result: Per-repository retention policies and audited deletion of pre-2021 files containing PII.

Government

The challenge: A mid-sized New England municipality stored PII across OneDrive, SharePoint, Exchange, SMB file shares, and SQL databases, with a lean IT team and no clear picture of who had access.

What Concentric AI uncovered: Five TB of data across major repositories, with automated classification and privacy detection providing visibility into sensitive information and access to it.

The result: A stronger data security foundation, including Gen AI DLP to monitor staff AI usage from day one without increasing headcount.

About Concentric AI

Concentric AI enables organizations to accelerate safe AI adoption through its unified AI and data security governance platform. Powered by the Semantic Intelligence Engine's patented language models, the platform discovers, governs, and protects business-critical and compliance-sensitive data at rest, in motion, and used by AI through a consistent policy framework.

Request a Demo