

Microsoft Copilot

The Challenge

A professional services organization wanted to deploy Microsoft Copilot across their workforce to boost productivity and provide employees with AI-driven assistance in everyday tasks.

The vision: empower teams to quickly access insights, automate workflows, and make decisions with AI support.

The obstacle: their data.

Copilot's responses were inconsistent. Outdated and irrelevant documents surfaced in answers. Sensitive HR and legal files appeared in searches intended for general users. Employees quickly lost trust in the system.

The enterprise's Microsoft 365 environment held over 39TB of unstructured data spread across SharePoint, OneDrive, Teams, and Exchange—full of duplicate files, ROT data (redundant, outdated, trivial), and no central taxonomy. Feeding this into Copilot risked inaccurate answers and compliance violations.

What started as a productivity initiative became a data governance challenge.

The Breakthrough

The company turned to Aparavi's Data Suite (ADS) and Data ToolChain (DTC) to prepare their Microsoft 365 environment for Copilot—**delivering AI-ready data pipelines in just days, not months**. Together, these tools rapidly transformed their unstructured sprawl into structured, intelligent catalogs.

With ADS, they:

- ✓ Connected and indexed data across SharePoint, OneDrive, Teams, and Exchange without migration.
- ✓ Normalized file formats and cleaned metadata across systems.
- ✓ Classified content by department, sensitivity, and business function.
- ✓ Created curated, role-specific datasets for Copilot to use—**ready within days**.

Then with DTC, they:

- ✓ Chunked and embedded each dataset for optimal LLM performance.
- ✓ Built tailored ingestion workflows to prevent cross-departmental data exposure.
- ✓ Delivered clean, permissions-aware data to Copilot for contextual, accurate responses.
- ✓ Enabled full traceability from query to source file for compliance oversight.

From unstructured sprawl to Copilot-ready in record time, ADS and DTC created a secure, scalable pipeline to fuel Microsoft 365 with clean, intelligent data.

Results at a glance

39TB of unstructured Microsoft data mapped & prepped

5 role-specific Copilot datasets curated and deployed in DAYS

18.4TB of outdated or duplicate data removed

2x improvement in Copilot response relevance and accuracy

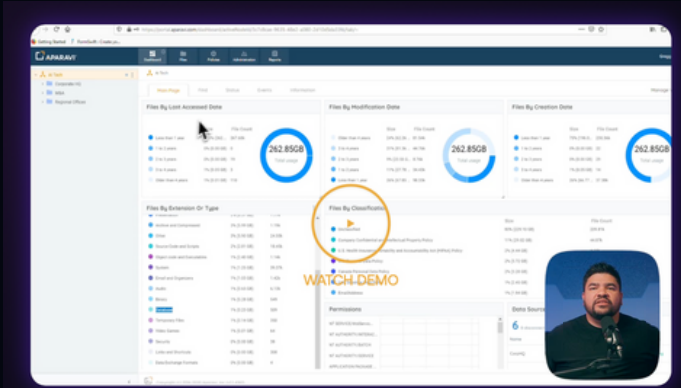
Zero unauthorized data exposures during Copilot queries

[Detailed view]

[Data Visibility, Control & Structure]

Aparavi Data Suite

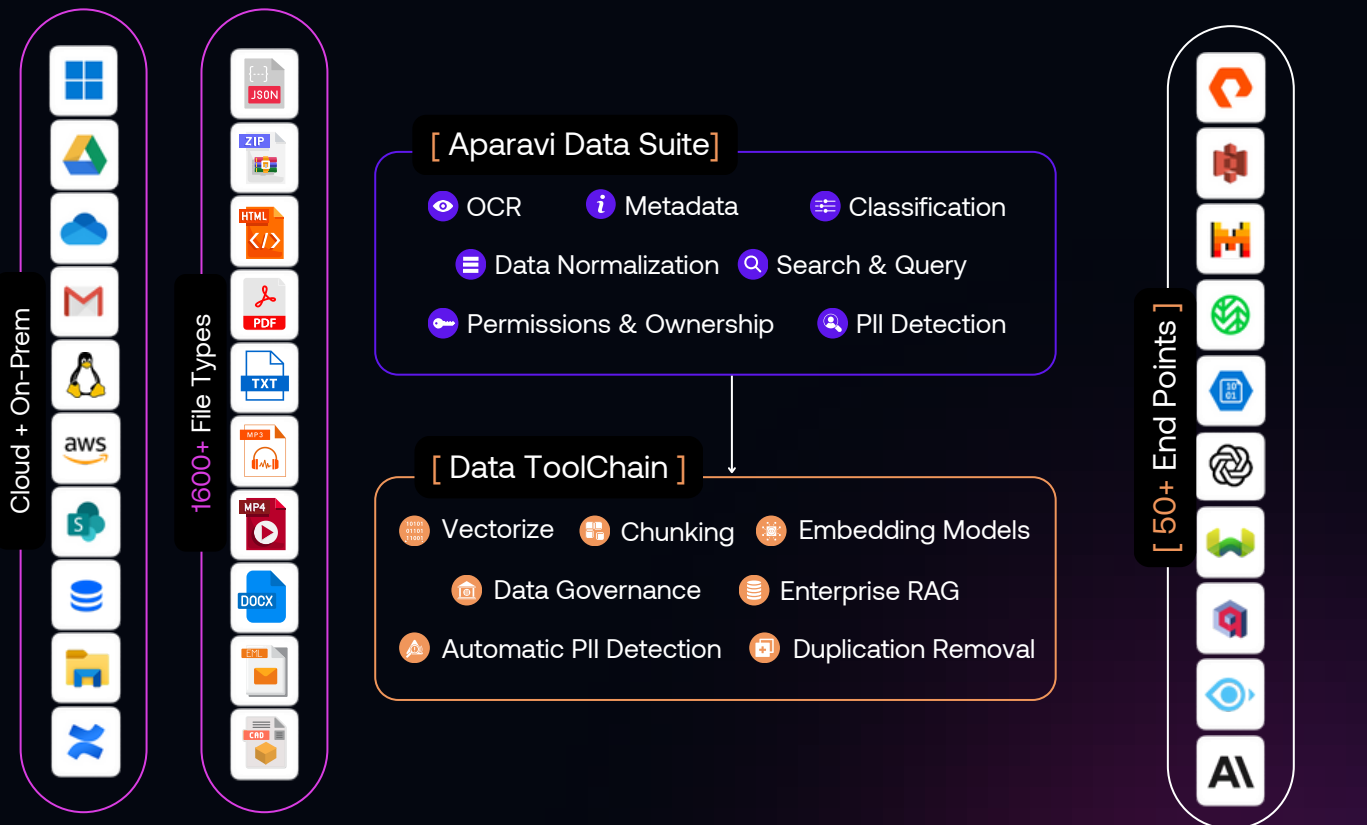
- Connected 39TB of Microsoft 365 data across environments—Without migrating anything, ADS securely indexed SharePoint, OneDrive, Teams, and Exchange to centralize visibility within hours.
- Normalized formats and standardized metadata—Cleaned inconsistent file naming, version sprawl, and malformed metadata for Copilot-ready content.
- Classified and tagged files by department, sensitivity, and business relevance—Using customizable policies to separate HR, finance, legal, and IT data.



[Data collection]

[Preparation & Workflows]

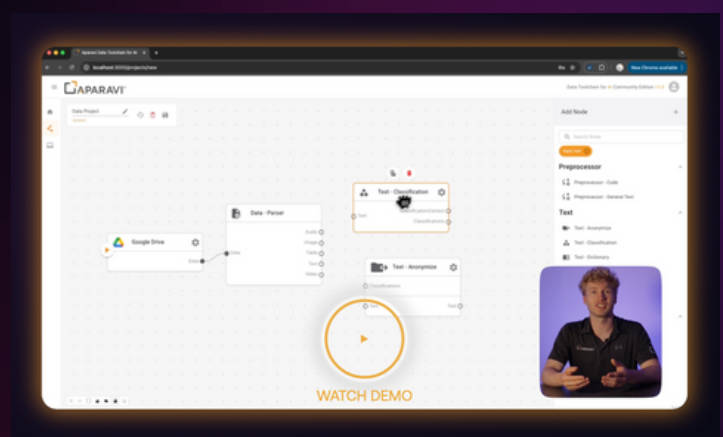
[Realization]



[Pipeline Execution, Vectorization & Deployment]

DTC

- Deployed Copilot-ready data with minimal IT lift—Using drag-and-drop workflows and ready-to-deploy formats, Copilot quickly began delivering value across departments
- Created a sustainable framework to scale Copilot use—The modular pipeline allows additional departments or business units to onboard Copilot seamlessly.
- Chunked and embedded curated datasets for optimized LLM performance—DTC automatically split data into context-aware, token-efficient chunks optimized for Copilot queries.
- Enabled full traceability from user query to source file—Critical for regulated industries to maintain confidence and compliance.



[Benefits]

Smarter Copilot Starts with Smarter Data

Aparavi helps enterprises rapidly prepare Microsoft environments for Copilot by transforming massive volumes of scattered, unstructured data into clean, classified, and AI-ready knowledge—all in days, not months—ensuring Copilot delivers accurate, compliant, and context-aware answers.



Context-Aware Copilot

Classified, curated data ensures Copilot retrieves only relevant and role-appropriate information—eliminating hallucinations and unauthorized exposure.



Governed from Day One

Automate retention and access policies to align with corporate and global regulations like ISO, NIST, and GDPR—so Copilot stays compliant as it learns.



A Repeatable Pipeline

From data discovery, vectorization and deployment, the same pipeline supports future Copilot expansions—reducing complexity and enabling scalable governance.

Spend less time sorting through messy data and more time making it work for you.

[Start Today](#)