Unstructured AI

Turn unstructured content into governed, AI-ready assets.

Enterprises generate vast amounts of unstructured content—documents, conversations, contracts, emails, transcripts and more. Yet most of it remains **locked away, fragmented and invisible** to AI systems. This blind spot delays projects, degrades accuracy and prevents teams from realizing AI's full potential.

Unstructured AI transforms this challenge into an opportunity by making unstructured data **discoverable, enriched and searchable at scale.** With advanced enrichment and automation, teams can rapidly extract business value, reduce manual work and fuel trustworthy, high-impact AI.

Benefits

Boost data usability
Expand the share of enterprise data usable for AI from a small fraction to a meaningful percentage by enriching and tagging unstructured content with high-fidelity metadata.

Improve AI and search accuracy
Prevent knowledge search degradation by enriching enterprise search and AI systems with semantic context, boosting retrieval accuracy and model outcomes.

Accelerate time-to-insight
Replace weeks of manual labeling and taxonomy work with automated enrichment. Rapidly tag, filter and structure content, reducing time from months to minutes.

Solution highlights

Smart discovery

Leverage an automated workflow to scan large repositories and identify relevant data for each use case with semantic tagging.

Automated semantic layer

Replace manual labeling and over-reliance on domain experts with auto-generated taxonomies that are learnt from your data.
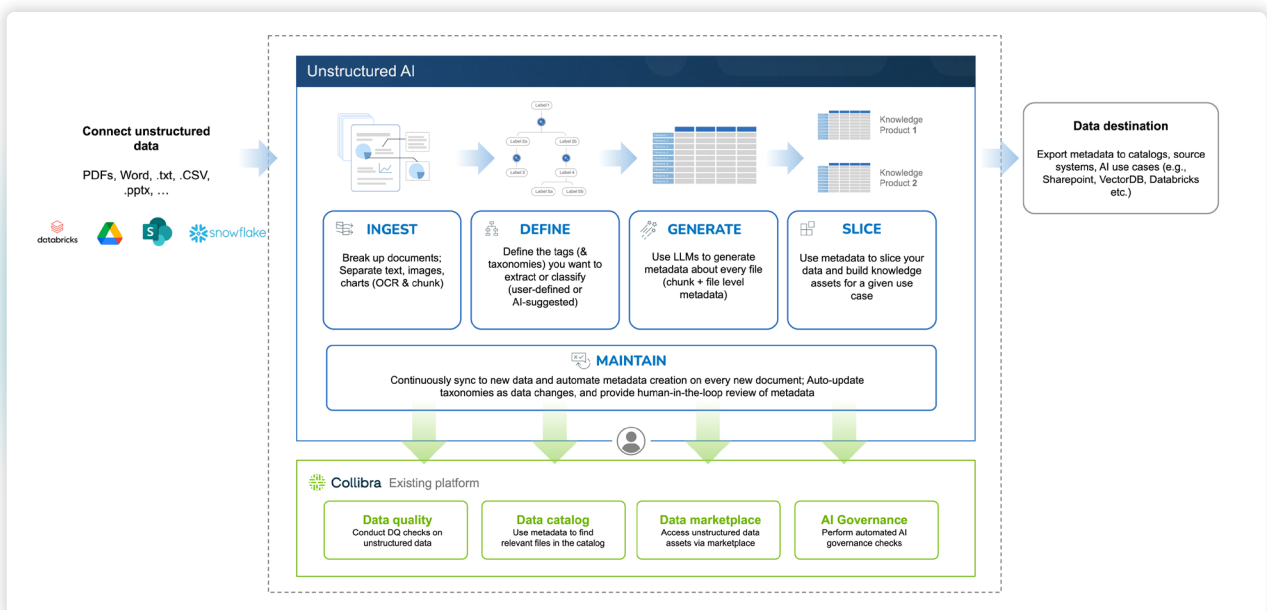
High-accuracy enterprise AI search

Maintain precision and context across thousands of files by enriching search and AI tools with semantic metadata.

High-fidelity metadata

Standardized, multimodal and high-quality metadata ensures effective classification and recall for enterprise-scale knowledge discovery.

Human-in-the loop visibility

Retain confidence in your metadata through clear confidence scoring, traceability and user-friendly fine-tuning workflows.

Automated metadata system

Rely on a system that continually tags new data produced in the enterprise and dynamically updates taxonomies as content evolves over time.

Proven business impact

- **3%-20%** Increase in total enterprise data usable for AI with a semantic tagging layer
- **1 month -20 min** Time saved to tag, organize and filter a corpus of 10,000 files
- **78% -92%** Uplift in search accuracy when adding semantic metadata into enterprise search systems

Fully automates the tagging, filtering and enrichment of unstructured data for AI

Want to see Collibra Unstructured AI in action?
Take the product tour at collibra.com.