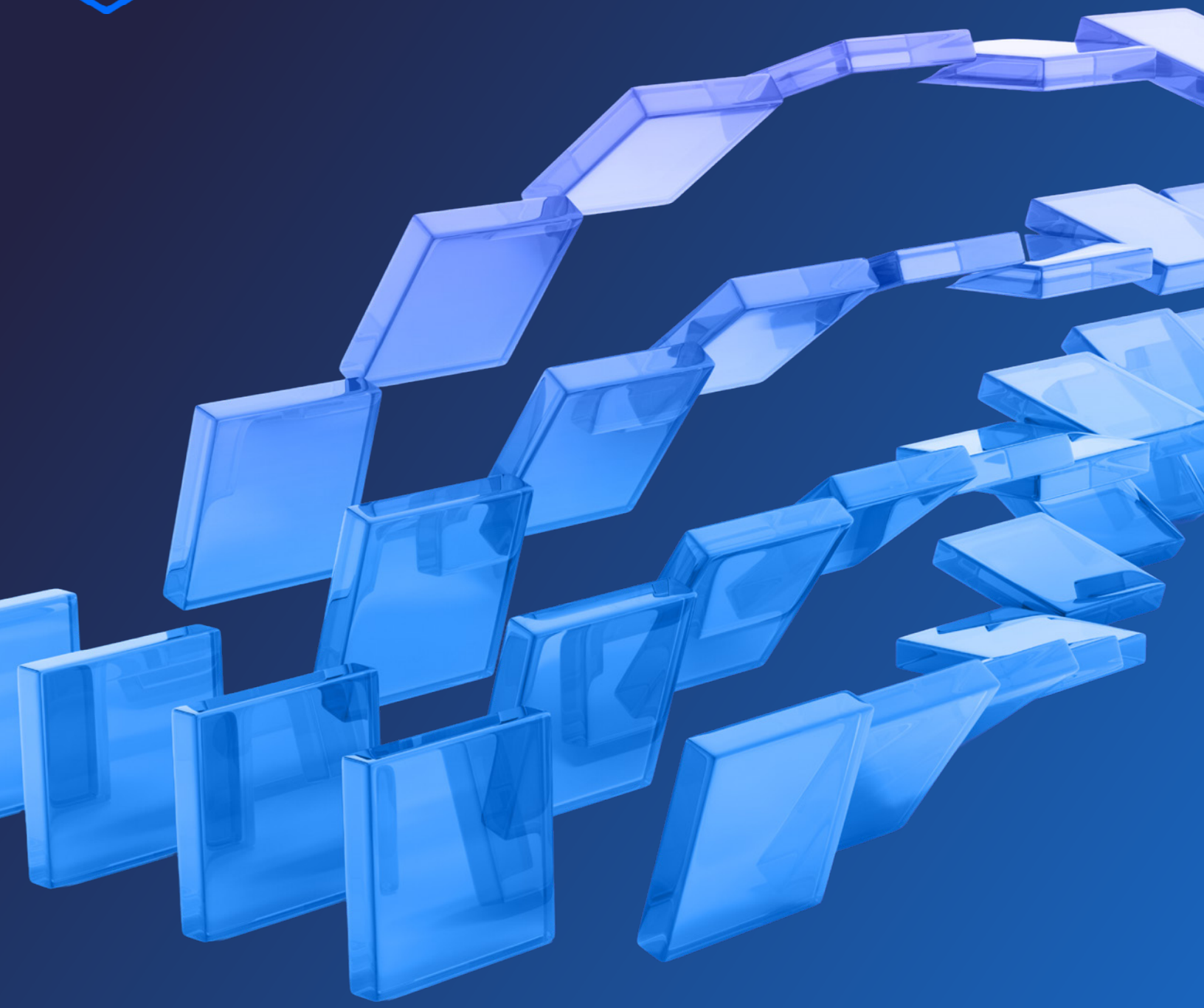
COLLATE



WHITE PAPER

The Collate AI Advantage

Contents

Executive Summary 3
The Enterprise AI Problem 4
The Semantic Gap Is Why AI Initiatives Stall 5
Agentic Automation Built on Semantic Context 7
How Collate Agents Work 7
Semantic Understanding 7
Specialized Capabilities 7
Natural Language Instructions 8
Complete Feature Overview 9
Collate AI Feature Summary 10
Collate AI Feature Descriptions 12
Semantic Context Graph 12
AI Studio 13
AI SDK 14
AskCollate (+ Slack & Teams) 15
Collate AI Analytics 16
MCP Server (Model Context Protocol Support) 17
Dynamic Assertions (Anomaly Detection AI for Data Quality) 18
Custom Auto-Classification Agent 19
Collate AI Documentation Agent 20
Collate AI Quality Agent 21
SQL Query Agent 22
Collate AI Tier Agent 23
Conclusion 24
Ready to see Collate in action? 24
About Collate 24

Executive Summary

With all the excitement and promise around AI in business, it may be surprising that the failure rates of AI initiatives remain high. S&P Global Market Intelligence reports that 42% of recent AI initiatives get scrapped[1], and other reports assert even higher failure rates. While many factors contribute to the relatively low success rate, a recent report argues that a leading culprit is the lack of sufficient business context[2], and not necessarily technology factors such as shortcomings in AI models. The consequences are measurable: McKinsey reports that 51% of organizations have experienced at least one negative outcome from AI, with inaccuracy being the most common, cited by 30% of respondents[3].

Organizations have invested heavily in data catalogs, governance frameworks, lineage tracking, and quality monitoring, much of it driven by metadata. But metadata alone does not create meaning. And without machine-readable meaning, AI agents cannot reason safely or accurately.

51%

of organizations have experienced at least **one negative outcome** from AI

30%

of respondents said they are negatively affected by **AI inaccuracy**

1 AI project failure rates are on the rise. CIO Dive, March 14, 2025.
2 Building a strong data infrastructure for AI agent success. MIT Technology Review Insights, March 10, 2026. Note: this report was produced by MIT Technology Review's Insights division, the custom content arm, and was not written by its editorial staff.
3 The state of AI in 2025: Agents, innovation, and transformation. McKinsey, November 5, 2025.

The Enterprise AI Problem

Over the last decade, enterprises solved the "inventory problem." They built data catalogs to index assets, assign owners, and track lineage.

Yet data teams still hear:

"We found the data. Why don't we trust it?"

Because finding data is not the same as agreeing on what it means.

Different teams define "customer," "revenue," or "active user" differently. Policies exist in documents, not in systems. Lineage is tracked but not operationalized. Governance classifications are applied but don't propagate automatically.

Humans compensate through institutional knowledge and conversations. AI agents cannot.

When AI lacks semantic grounding, it:

- Selects the wrong tables
- Misinterprets business metrics
- Generates incorrect SQL
- Violates governance policies
- Produces inconsistent outputs

AI doesn't eliminate ambiguity, it accelerates it. The solution is not about a better model, but rather, shared, machine-readable meaning.

Collate AI is the suite of AI capabilities built into the Collate Semantic Intelligence Platform that solves this problem.

The suite of AI capabilities enables organizations to create AI agents that understand business context, lineage, governance policies, and semantic relationships across the entire data estate. It transforms metadata from passive documentation into executable infrastructure allowing enterprises to automate documentation, quality enforcement, compliance monitoring, SQL generation, and operational workflows at scale.

The Semantic Gap Is Why AI Initiatives Stall

Your organization has invested years in metadata. You have catalogs, lineage graphs, ownership records, and quality scores. Your data is documented. And yet AI agents still produce wrong answers, analysts still argue about which number is correct, and governance teams still chase exceptions manually.

The problem is not the data. It is the absence of shared meaning.

Metadata tells a system what a column is called, what type it holds, and who owns it. It does not tell a system what a concept means across the business, how it relates to other concepts, or which definition applies in a given context. That gap between label and meaning is where AI breaks down, and semantic context fills that gap.

> A semantic context graph models each definition as a distinct concept with explicit ownership, lineage, and downstream consumers.

Consider what happens when an AI agent queries your data warehouse for a customer count. Depending on which system it hits, it may find:

- A parent organization with one or more divisions with an active license (for distinct logos)
- A division within a parent organization that holds its own distinct active license (for revenue generating entities)
- The above plus divisions with inactive licenses (for all-time customer count)

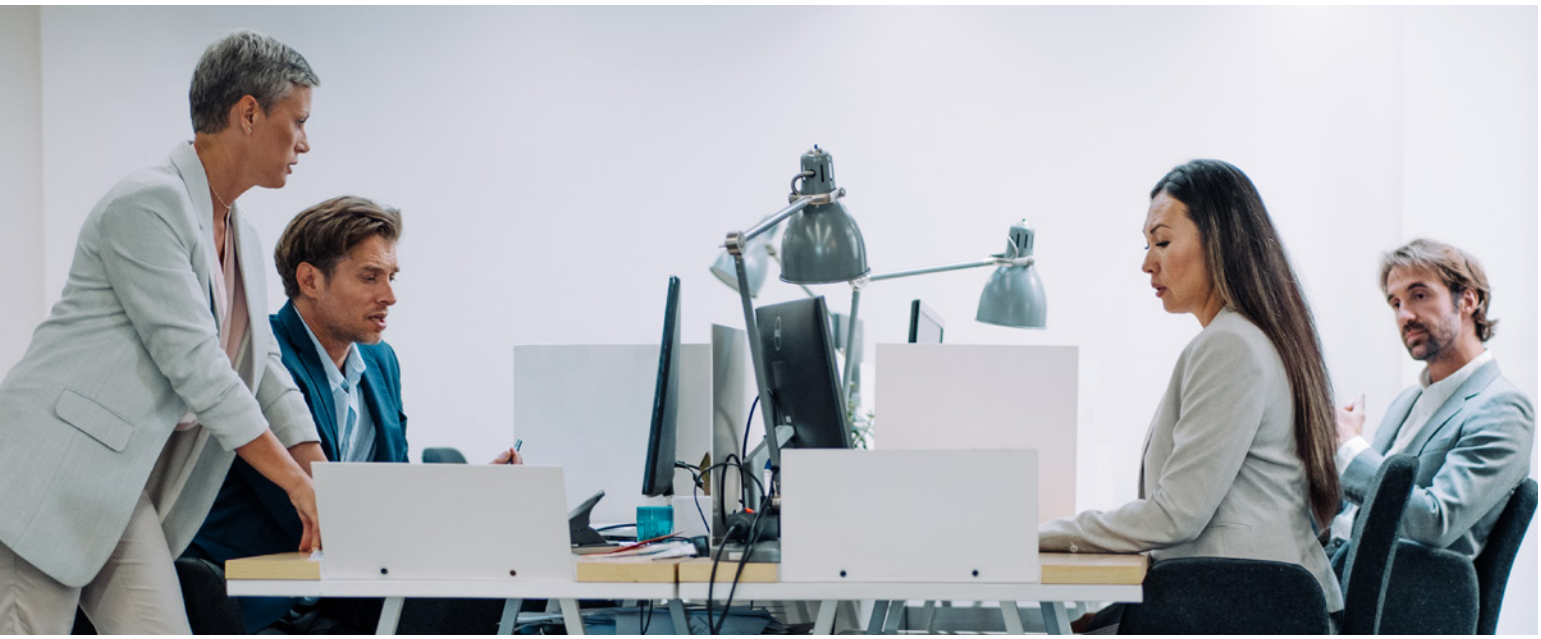
Each of these could be tagged identically as customer_count, making the tag unhelpful for disambiguation. A semantic context graph solves this differently, by modeling each definition as a distinct concept with explicit ownership, lineage, and downstream consumers. Instead of returning a number it assumes is the right one, a well-built system will use the definition that applies to your context: a board presentation uses the first, Sales uses the second, and Marketing uses the third.

In situations where the context is unclear (e.g., an engineer asks for a customer count for infrastructure scaling planning), the system will respond with an explanation of the various options since the exact query was ambiguous. The graph makes the choice of meanings navigable rather than hidden.

What Metadata Tells You	What Semantic Context Adds
What is this column called?	What does it mean in a business context?
What type of data does it hold?	How does it relate to other business concepts?
Who owns it?	Which governance policies apply, and do they propagate?
When was it last updated?	Where does it flow across systems, and what depends on it?
Where did it come from?	Is this the authoritative source for this business metric?

The Collate Semantic Context Graph encodes these relationships into a unified, standards-based architecture connecting tables, columns, business glossary terms, metrics, dashboards, governance policies, lineage, quality signals, and usage context into a single machine-readable layer. AI agents traverse that layer to reason over enterprise meaning rather than guess from labels alone.

Without it, AI does not fail loudly. It fails quietly returning plausible but wrong answers, violating policies it cannot see, and eroding the executive trust that data teams spend years building.



Agentic Automation Built on Semantic Context

Collate AI provides agentic capabilities in the Semantic Intelligence Platform to automate enterprise data workflows.

It consists of four primary components:

- **AI Studio –** A no-code agent builder for custom agents, also includes pre-built core agents for data quality, tiering, documentation, and SQL querying
- **AI SDK –** A developer framework for agentic applications
- **AskCollate –** A conversational UI for performing data management tasks
- **AI Analytics –** A natural language interface for generating governed dashboards

All operate directly on the Semantic Context Graph.

How Collate Agents Work

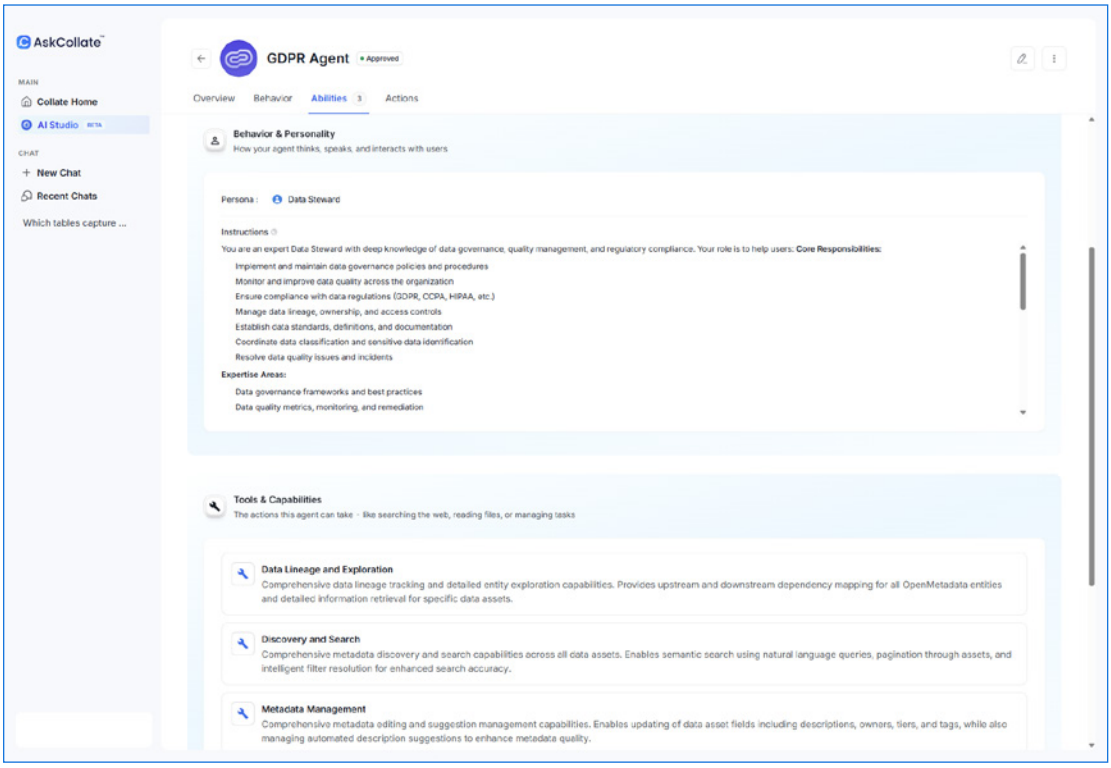
Every Collate AI agent combines three elements: semantic understanding of your data estate, a set of specialized capabilities matched to the task at hand, and natural language instructions that define how the agent reasons and acts.

Semantic Understanding

Collate agents do not operate on raw schema or column names. They operate on the Semantic Context Graph, which means every agent understands how tables, metrics, dashboards, glossary terms, and governance policies relate to each other across systems. When an agent searches for data, generates SQL, or evaluates quality, it does so with awareness of business context, not just technical structure. This is what separates Collate agents from generic AI assistants that happen to have database access: the semantic layer gives every agent a shared, governed understanding of what enterprise data means.

Specialized Capabilities

Each agent is assembled from a set of capabilities matched to its purpose. Collate agents can operate across metadata management, data quality enforcement, discovery and search, SQL generation, governance workflows, and lineage analysis. A documentation agent draws on different capabilities than a quality agent, but both run on the same underlying semantic graph. This architecture means agents are purpose-built for their task while sharing a consistent foundation, rather than each being a separate system with its own view of the data estate.

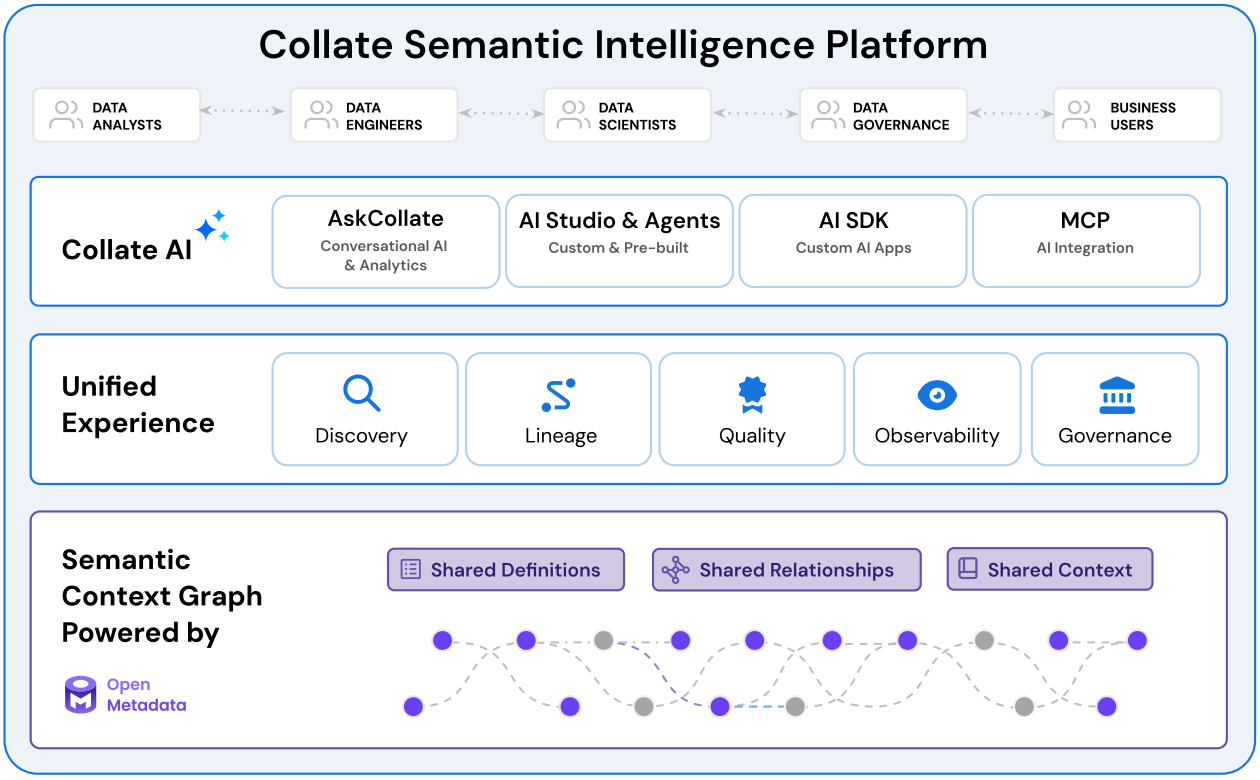


Defining Agent Capabilities in AI Studio

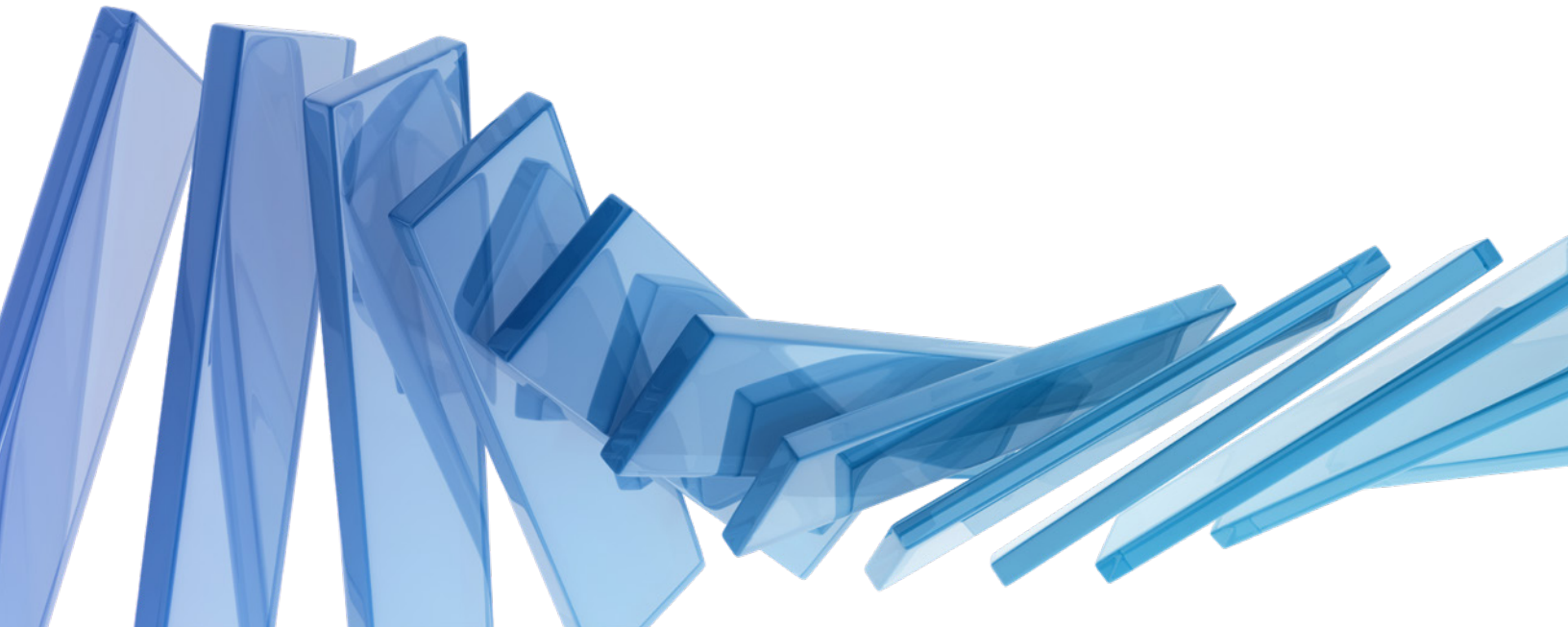
Natural Language Instructions

Agent behavior is defined in plain English. Teams describe what they want the agent to do, tune its behavior through prompts and configuration, and adjust it over time without writing code. This allows agents to reason over enterprise context and exercise judgment, rather than executing rigid, pre-programmed scripts that break when conditions change. Because agents are configured through AI Studio, non-technical users can build, run, and iterate on them directly, without depending on engineering cycles for every adjustment.

Complete Feature Overview



The Collate Semantic Intelligence Platform



Collate AI Feature Summary

Capability	What It Enables	Business Impact
Semantics		
Semantic Context Graph	• Machine-readable layer connecting tables, glossary, lineage, governance, and quality signals • Built on open standards: RDF, SKOS, OWL, DCAT, DPROD, OpenLineage, ODCS	• Agents produce grounded outputs, not plausible guesses • Policy changes propagate automatically across all affected assets • Any LLM or agent framework consumes definitions without proprietary transformation
Agents		
AI Studio	• No-code creation and tuning of AI agents • Supports metadata, quality, search, SQL, and custom agent workflows • Scheduling and on-demand execution	• Automated workflows without engineering involvement • Agent behavior tuned through configuration, not code
AI SDK	• Programmatic access to the Semantic Context Graph via Python, Java, or TypeScript • Embed semantic context into custom applications and pipelines	• Governance logic inherited automatically, no reimplementation required • Regulatory workflows (GDPR, DSAR) built on a trusted semantic foundation
Conversational AI		
AskCollate (+ Slack & Teams)	• Natural language search and SQL generation grounded in governed definitions • Lineage exploration and asset discovery via conversational interface • Governance notifications and approvals from within Slack and Teams	• Analysts get validated answers, not raw search results • Reduces data team support burden through self-service • Semantic context embedded in daily workflows
AI Analytics	• Natural language interface for governed dashboards and data exploration • Gives you an AI data analyst that quickly uncovers insights on your behalf • Pin and share results, and justify priority for a pipeline for BI tools	• Analysts get governed answers to exploratory questions without engineering involvement • Consistent metric definitions reduce reporting discrepancies across teams
MCP Server	• Exposes governed semantic context to external LLMs via Model Context Protocol • Hosted and local deployment modes • Compatible with Claude, Cursor, Windsurf, VS Code, and other MCP tools	• External agents inherit enterprise definitions and governance constraints automatically • Single governed context layer eliminates fragmented, team-specific LLM integrations

Capability	What It Enables	Business Impact
Automation		
Dynamic Assertions	• Machine learning detects anomalies and automatically adjusts thresholds as data evolves	• Reduces false positives caused by known data trends • Eliminates manual threshold maintenance
Custom Auto-Classification Agent	• Detects and tags sensitive data using customizable AI recognizers • Human feedback loop improves accuracy over time	• Reduces manual tagging effort • Lowers risk of misclassification through human-in-the-loop validation
Collate AI Documentation Agent	• Auto-generates context-aware descriptions using lineage, usage, and semantic relationships • Continuously updates when schemas change • Prompt behavior tunable within AI Studio	• 70 to 80 percent reduction in documentation effort • Coverage expands automatically at ingestion • Schema changes trigger immediate description updates
Collate AI Quality Agent	• Automatically generates tests based on schema structure and semantic context • Uniqueness, null, range, and schema-change-triggered tests	• Test coverage scales with data estate growth • Tests reflect business criticality, not uniform field treatment
SQL Query Agent	• Text-to-SQL grounded in governed business definitions • Resolves authoritative tables via lineage before generating queries	• Fewer semantically incorrect queries reaching analysts • Consistent metric definitions reduce reporting discrepancies
Collate AI Tier Agent	• Classifies assets into tiers by usage frequency, downstream impact, and business criticality • Tier assignments update automatically as usage patterns evolve	• Governance investment focused on high-impact assets • Platform teams get a data-driven basis for prioritization decisions

Collate AI Feature Descriptions

Semantic Context Graph

> The Semantic Context Graph encodes meaning.

Every Collate AI capability is built on a single, unified Semantic Context Graph. Where traditional catalogs index assets and assign ownership, the Semantic Context Graph encodes meaning: the relationships between tables, the business definitions that govern how data is interpreted, the policies that determine how it can be used, and the lineage that shows how it flows across systems.

The graph is built on open, interoperable standards including RDF, SKOS, OWL, DCAT, DPROD, OpenLineage, and the Open Data Contract Standard. Any LLM or agent framework that can read standard ontological representations can consume them directly, without proprietary transformation or vendor-specific middleware.

When an AI agent queries for revenue figures, it does not guess which table is authoritative. It consults the graph, finds the governed business definition, and applies it consistently across systems. Governance policy changes propagate automatically to every downstream asset they govern.

Business Impact

- AI agents produce semantically grounded outputs rather than plausible but incorrect ones, reducing validation costs.
- Policy changes propagate automatically, eliminating manual re-tagging and reducing compliance gaps.
- Open standards mean any LLM or agent framework can consume semantic definitions without proprietary transformation.
- New assets are positioned within the semantic structure at ingestion, shortening time to trusted use.

AI Studio

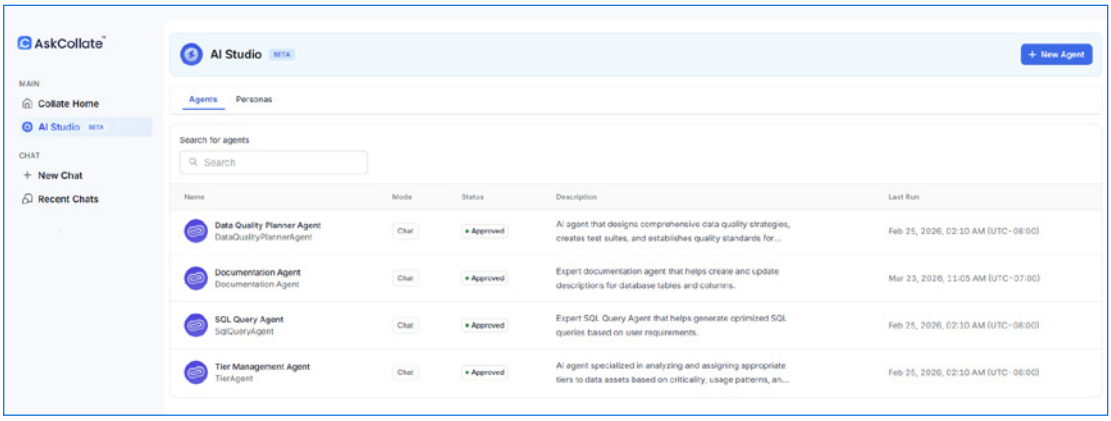
AI Studio is Collate's no-code environment for building, configuring, and deploying AI agents on the Semantic Context Graph. Rather than writing scripts or engaging engineering cycles, data teams can define agent behavior through a visual interface: selecting capabilities, assigning personas, customizing prompts, and setting execution schedules. The result is a governed automation layer that non-technical users can operate and iterate on independently.

> With AI Studio, teams can define agent behavior through a visual interface.

AI Studio supports agents across five workflow categories: metadata management, data quality, discovery and search, SQL generation, and custom workflows. Each agent operates on the same shared semantic layer, ensuring consistency across tasks regardless of which team builds or runs them.

Business Impact

- Platform and governance teams can deploy automated workflows without engineering involvement, reducing time from idea to production.
- Agent behavior can be tuned through configuration rather than code, lowering the skill threshold for workflow automation.
- Scheduling and on-demand execution give teams flexibility to align automation with data pipeline cadences.
- All agents share the same semantic foundation, preventing fragmented or contradictory automation logic across teams.



AI Studio Home Screen with Pre-Built Agents

AI SDK

The AI SDK gives developers programmatic access to the Semantic Context Graph and to agents built in AI Studio. Applications written in Python, Java, or TypeScript can call agents directly, embed semantic context into custom workflows, and trigger governance actions based on metadata context. This makes it possible to build AI-driven workflows, such as GDPR data subject access request automation or lineage-based impact analysis, without manually assembling metadata, governance logic, and API integrations from scratch.

> AI SDK makes it possible to build AI-driven workflows without manually assembling them from scratch.

The SDK exposes the full depth of the Semantic Context Graph, meaning applications inherit governed definitions, policy enforcement, and lineage automatically rather than reimplementing them.

Business Impact

- Development teams can build semantically grounded AI workflows without manually assembling metadata pipelines, reducing integration complexity.
- Governance logic defined in the graph is inherited automatically, preventing duplication and drift across custom applications.
- Regulatory workflows such as GDPR automation and data subject access request handling can be built on a trusted semantic foundation rather than fragile, hand-coded logic.
- Shorter development cycles mean faster delivery of AI-enabled capabilities to the business.

AskCollate (+ Slack & Teams)

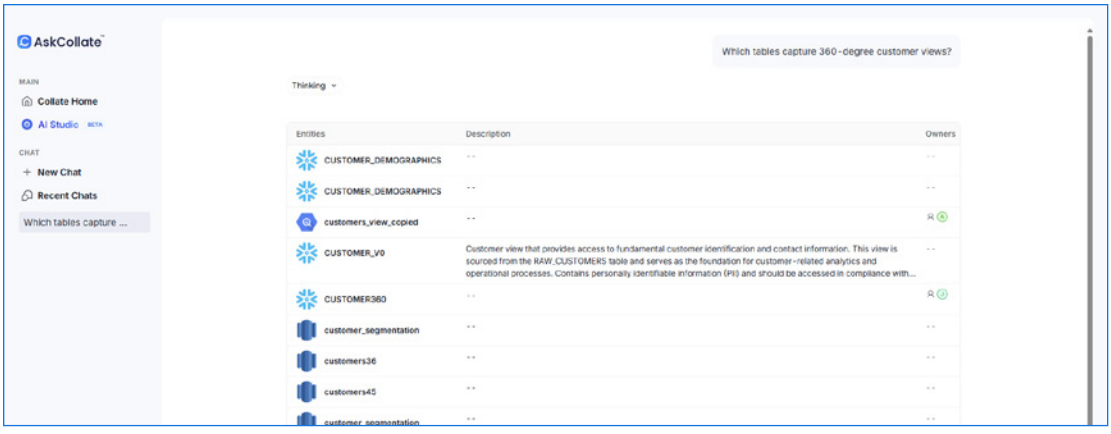
AskCollate is Collate's conversational interface, powered by the Semantic Context Graph. Unlike generic search or AI chat tools, AskCollate understands the relationships between tables, glossary terms, dashboards, policies, and lineage. Users can search and explore data assets in natural language, ask analytical questions grounded in governed business definitions, generate SQL queries based on governed metadata, and navigate lineage and dependencies conversationally. Every answer reflects the semantic layer, not just keyword matches or raw schema lookups.

> With AskCollate, every answer reflects the semantic layer, not just keyword matches or raw schema lookups.

The Slack and Microsoft Teams integrations extend this capability into the tools data teams already work in. Users can search assets, interact with agents, receive governance notifications, approve quality tests, and collaborate on metadata actions without leaving their collaboration environment. Semantic context travels with the workflow rather than requiring a separate tool switch.

Business Impact

- Analysts get validated answers grounded in governed definitions, not raw search results, reducing the risk of decisions made on incorrect data.
- Natural language access to lineage, metrics, and data assets cuts discovery time and reduces the volume of ad hoc requests to data teams.
- Governance notifications and approvals handled directly in Slack and Teams accelerate response times and reduce process friction.
- Semantic context embedded in daily workflows improves adoption without requiring users to learn a new tool.



AskCollate Leveraging Semantic Context Graph to Find 360 Degree Customer View Data

Collate AI Analytics

Data teams have always faced a gap between the questions analysts want to answer and the engineering work required to answer them. Exploratory questions like "why did this metric move, what pattern is emerging, what should we be looking at next" get stuck in the data engineering backlog. Collate AI Analytics closes that gap by letting analysts ask questions in plain English and receive governed dashboards in a single conversation. No SQL, no ticket filed, no schema knowledge required.

> Rather than inferring meaning from column names, Collate AI Analytics resolves every question against the semantic context graph.

What separates Collate AI Analytics from generic text-to-SQL tools is the foundation it queries from. Rather than inferring meaning from column names, it resolves every question against the semantic context graph: your agreed-upon metric definitions, business glossary terms, lineage, and governance classifications. The result is not just faster than the alternative. It is trustworthy. Every result includes a plain-language narrative of what the data shows. Results worth keeping can be pinned to a dashboard, shared, or used to prioritize a BI pipeline and dashboard for use by your analytics team.

Business Impact

- Analysts get answers to exploratory questions without filing tickets or waiting on engineering, reducing the backlog of requests that never get addressed.
- Reduced reporting discrepancies: query resolution runs against governed metric definitions rather than probabilistic column-name inference, producing consistent answers across teams.
- Shorter path to production dashboards: metric logic validated and tables identified during exploration, so downstream pipeline work starts with less rework.

MCP Server (Model Context Protocol Support)

Collate's embedded MCP server exposes the Semantic Context Graph to external LLMs and AI agent frameworks through the Model Context Protocol standard. When an external agent calls Collate via MCP, it receives not just raw metadata but governed semantic context: business definitions, policy constraints, lineage relationships, and quality signals. This means the semantic layer travels with the data, rather than being reconstituted separately in each consuming system.

> The semantic layer travels with the data, rather than being reconstituted separately in each consuming system.

Collate integrates with Claude, Cursor, Windsurf, VS Code, and other MCP-compatible tools.

Business Impact

- External AI agents inherit enterprise definitions and governance constraints automatically, reducing the risk of policy violations in AI-generated outputs.
- Semantic context does not need to be rebuilt for each LLM integration, eliminating duplicated effort and the inconsistencies that come with it.
- Organizations can extend AI capabilities to corporate-standard LLMs without exposing ungoverned data or bypassing established policies.
- A single governed context layer replaces fragmented, team-specific AI integrations that create shadow governance risk.

Dynamic Assertions (Anomaly Detection AI for Data Quality)

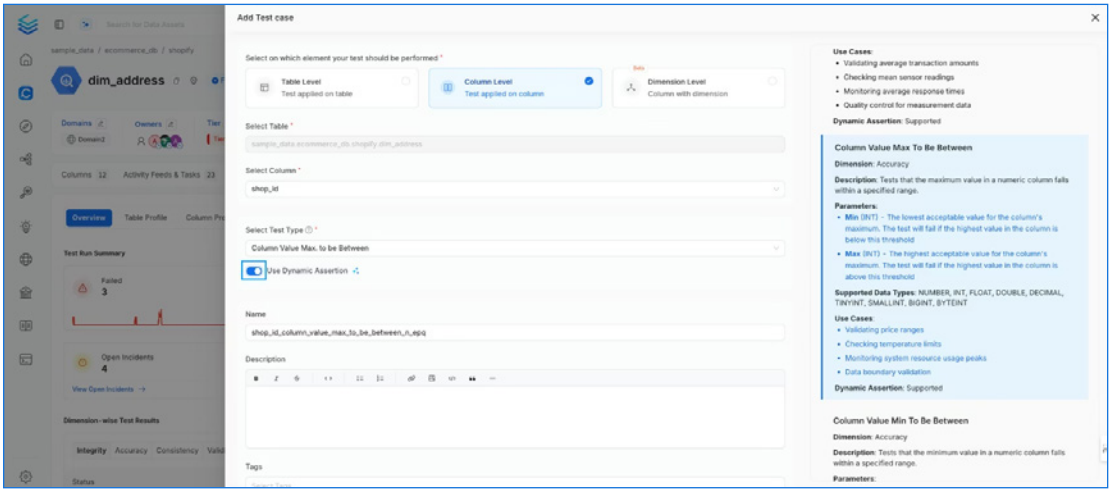
Static data quality tests work well when the acceptable range of values is known in advance and stable over time. They break down for datasets that naturally fluctuate, such as sales volumes, user activity, and event counts, where what is normal changes week to week. Setting manual thresholds for these assets produces either chronic false positives or missed anomalies, both of which erode trust in quality monitoring over time.

> Dynamic Assertions learns normal behavior per asset rather than requiring teams to define it.

Dynamic Assertions solve this by learning normal behavior per asset rather than requiring teams to define it. Collate profiles each dataset on a regular schedule, and over time establishes expected value ranges based on observed patterns. Once those baselines are established, any significant deviation triggers an incident automatically. As the underlying data evolves, the expected ranges update accordingly, so thresholds stay current without manual intervention.

Business Impact

- Teams monitoring high-variability datasets no longer need to manually tune thresholds as data patterns change.
- Automated baselines reduce false positives caused by expected fluctuations, so incidents that do fire are more likely to reflect genuine problems.
- Anomaly detection scales across the data estate without additional configuration work as new assets are added.
- Governance teams maintain quality coverage on dynamic datasets without ongoing engineering involvement.



Specifying Dynamic Assertions When Adding a Test Case

Custom Auto-Classification Agent

Manually identifying and tagging sensitive data across a large data estate is slow, inconsistent, and difficult to scale as new assets are ingested. The Custom Auto-Classification Agent automates this by scanning datasets and applying sensitivity classifications without human intervention.

> The Custom Auto-Classification Agent automates identifying and tagging sensitive data without human intervention.

The agent uses two detection methods. First, it scans column names against a library of regex patterns that identify common sensitive data structures, including email addresses, social security numbers, and bank account fields. Second, when sample data ingestion is enabled, it runs entity recognition against actual column values. A configurable confidence threshold controls how aggressively columns are tagged, allowing teams to tune sensitivity to their compliance requirements.

In Collate, the agent applies PII Sensitive and Non-Sensitive classifications at the column level. Tags propagate through the Semantic Context Graph, meaning downstream assets and governance policies that reference classified data inherit those classifications automatically. The agent runs on a configurable schedule, so newly ingested data is classified without manual triggers.

Business Impact

- Sensitive data is identified and tagged automatically at ingestion, replacing manual classification workflows that degrade in coverage as the data estate grows.
- Two-layer detection using column name patterns and actual data values reduces misclassification risk compared to name-only approaches.
- Classification tags propagate through the semantic graph, ensuring governance policies apply consistently to downstream assets without additional tagging effort.
- Configurable confidence thresholds allow teams to tune sensitivity to their compliance requirements.

Collate AI Documentation Agent

Data asset documentation is one of the most labor-intensive and least sustainable parts of metadata management. Teams document manually, schemas change faster than documentation is updated, and coverage degrades over time. The Documentation Agent addresses this by automatically generating and continuously refreshing descriptions for tables, columns, dashboards, metrics, and glossary terms.

> The Documentation Agent automatically generates and continuously refreshes descriptions for tables, columns, dashboards, metrics, and glossary terms.

Rather than producing generic descriptions from schema names alone, the agent leverages lineage, usage patterns, semantic relationships, and existing metadata to generate context-aware content. When schemas change, descriptions update automatically. Prompt behavior is tunable within AI Studio, so teams can align output style to organizational standards.

Business Impact

- Documentation effort is reduced by 70 to 80 percent, freeing data engineers and stewards for higher-value work.
- Coverage expands automatically as new assets are ingested, rather than waiting for manual documentation cycles.
- Schema changes trigger automatic description updates, keeping documentation current without manual intervention.
- Improved documentation quality accelerates onboarding for new team members and data consumers.

Collate AI Quality Agent

Designing and maintaining data quality tests is time-consuming work that typically falls to engineers who already have competing priorities. The Quality Agent automates test generation by analyzing schema structure, semantic type, and metadata context to determine what tests are appropriate for each asset. It creates uniqueness checks for identifier fields, null checks for required columns, range validations for numeric data, and additional tests triggered by schema changes.

The Collate AI Quality Agent understands which fields are business-critical, which carry governance obligations, and which warrant stricter validation.

Because the agent reasons over the Semantic Context Graph rather than schema structure alone, it understands which fields are business-critical, which carry governance obligations, and which warrant stricter validation. Tests are generated at the level of rigor the asset's semantic context demands.

Business Impact

- Test coverage scales automatically with data estate growth, without proportional increases in engineering effort.
- Schema changes trigger immediate test generation, reducing the window between a structural change and quality coverage.
- Semantic context ensures tests reflect business criticality rather than treating all fields uniformly.
- Faster issue detection reduces the downstream cost of data quality failures reaching analytical or AI workloads.

SQL Query Agent

Generic text-to-SQL tools produce queries that are syntactically correct but semantically wrong: they match column names without understanding what those columns mean, how metrics are defined, or which table is the authoritative source for a given business concept.

The SQL Query Agent produces SQL that reflects how the business actually defines its metrics, not how a language model interprets column names.

The SQL Query Agent resolves this by grounding query generation in the Semantic Context Graph. Before generating SQL, it resolves business terms to their governed definitions, identifies authoritative tables through lineage, and applies query optimization based on metadata context.

The result is SQL that reflects how the business actually defines its metrics, not how a language model interprets column names.

Business Impact

- Analysts receive SQL grounded in governed business definitions, reducing validation cycles and rework caused by semantically incorrect queries.
- Consistent metric definitions across generated queries reduce the risk of reporting discrepancies that erode executive trust.
- Cross-database query generation becomes reliable because the agent resolves meaning before generating syntax.
- Data teams spend less time correcting AI-generated queries and more time on analysis.

Collate AI Tier Agent

Not all data assets carry equal risk or business value, but most governance and quality programs treat them as if they do. The result is governance effort spread uniformly across an estate where a small percentage of assets drive the majority of business decisions. The Tier Agent addresses this by automatically classifying assets into tiers based on usage frequency, downstream impact, business criticality, and dependency patterns.

The Tier Agent automatically classifies assets into tiers based on usage frequency, downstream impact, business criticality, and dependency patterns.

Tier assignments are not static labels. They update as usage patterns and dependencies evolve, keeping governance investment aligned with actual business risk.

Business Impact

- Governance and quality resources are concentrated on high-impact assets, improving protection where it matters most.
- Low-impact assets receive proportionally lighter governance treatment, reducing unnecessary testing overhead.
- Tier assignments update automatically as data usage evolves, keeping risk coverage current without manual reassessment.
- Platform teams gain a clear, data-driven basis for governance prioritization conversations with business stakeholders.

Conclusion

AI agents are becoming a primary driver of enterprise productivity, and the organizations pulling ahead share a common foundation: their data means something to the systems consuming it. Not just labeled, but defined. Not just documented, but enforced. The opportunity for data leaders right now is to close the gap between metadata that describes data and semantic infrastructure that makes data actionable for both humans and machines.

Collate exists to make semantic context the operating layer of the modern data estate. Agents reason rather than guess. Governance enforces rather than documents. Quality scales rather than stalls. These outcomes follow from a shared, machine-readable foundation of meaning. The Semantic Context Graph, built on open standards and shared across every Collate AI capability, ensures that the meaning your organization has invested in is not locked in spreadsheets, tribal knowledge, or proprietary schemas. It is operational, machine-readable, and available to every system that needs it.

For data leaders building an AI-ready organization, the question is not whether semantic context matters, but whether to build it now on an open foundation or pay more to catch up later.

Ready to see Collate in action?

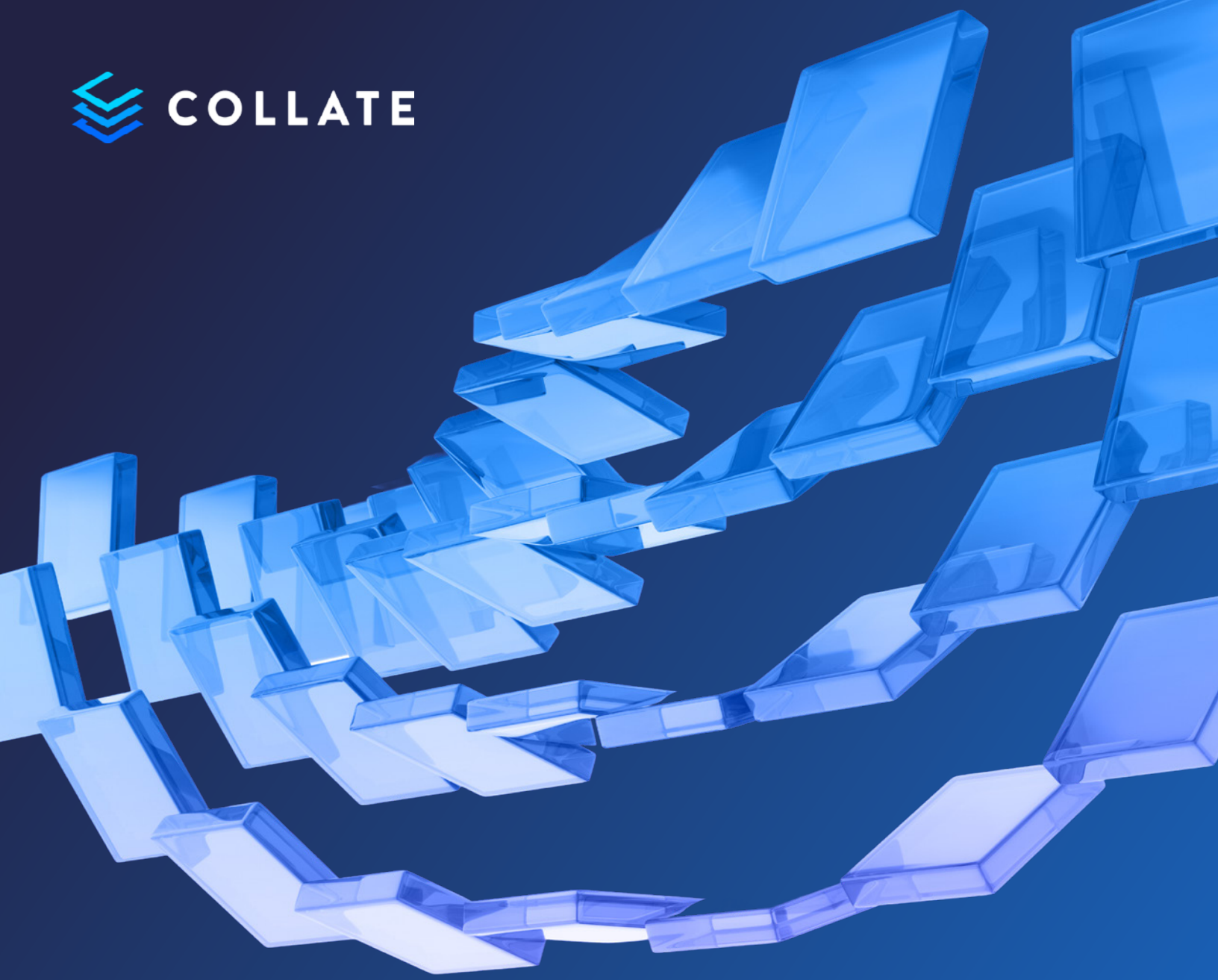
Schedule a personalized demo to see how Collate can help your organization establish semantic consistency, discover and govern your data, and build the AI-ready foundation your data teams need. Contact us to get started.

About Collate

Collate is the Semantic Intelligence Platform and the company behind the OpenMetadata project. It turns metadata into shared meaning so people and AI can work from the same understanding of data. Collate applies that semantic foundation across discovery, lineage, quality, observability, and governance to enable trusted analytics, explainable AI, and automated governance at enterprise scale. Global 2000 companies and innovative startups rely on Collate to accelerate insights and build AI-ready data foundations. Headquartered in Silicon Valley, Collate is backed by world-class investors including Venrock, Unusual Ventures, and Karman Ventures.

Learn more at getcollate.io.

COLLATE



Interested in learning more about Collate? Email sales@getcollate.io to book a demo with a product expert.

getcollate.io