← Back to blog

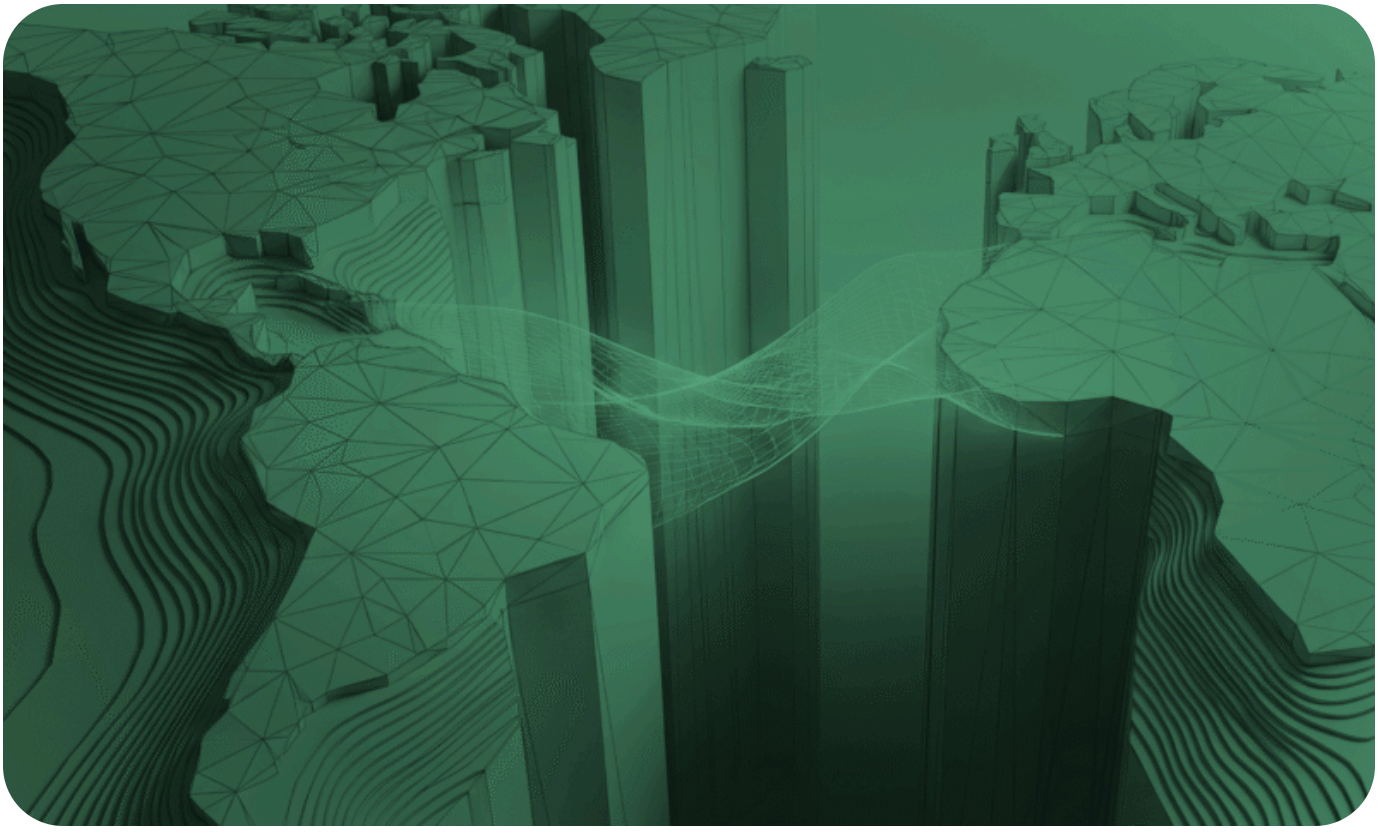
Sema4.ai v3.0, Powering Team-based Work and Lower Token Spend

Sema4.ai v3.0 advances the enterprise AI agent platform with shared workspaces, verified queries and logic, Data Collections, and ecosystem interoperability. Together, these capabilities help teams move complex, repeatable back-office work into governed, cost-efficient AI workflows.



October 29, 2026

Tim Rottach



Sema4.ai v3.0 is built to tackle complex, team-based work, with enhancements that help build trust in production agents and reduce operating costs.

The capabilities in the latest version of the Sema4.ai platform are expanded in the following areas:

- **Support for team-based, complex work:** teams share and manage agent work, skills, memories, data, and files within governed workspaces, for confidently collaborating across challenging team projects.
- **Ecosystem interoperability:** Organizations can now build and interact with Sema4 agents via AI tools like Claude Cowork and Microsoft CoPilot.
- **End-to-end lifecycle support:** Persuasive AI assistance is available in all aspects of the platform, from taking an agent from build to governing it, to tracking changes and audit reporting.
- **Streamlined automation and reduced token usage:** by automatically converting repetitive work to code, the platform reduces AI model calls, which improves accuracy and speed, while reducing cost and risk.

For teams responsible for complex, team-based work processes across finance, procurement, supply chain, logistics, and manufacturing operations, the **Sema4.ai Enterprise AI Agent Platform** helps turn repeatable, document- and data-heavy work into reliable workflows that can be shared, reviewed, optimized, and audited.

Complex business processes need something different

Most AI tools were designed to help an individual complete a task. Team-based work has additional challenges beyond those of the individual. AI agent platforms need a way for teams to share operational knowledge, connect documents and data, govern how agents run, reuse proven logic, manage exceptions, and trace outcomes back to the decisions and information behind them.

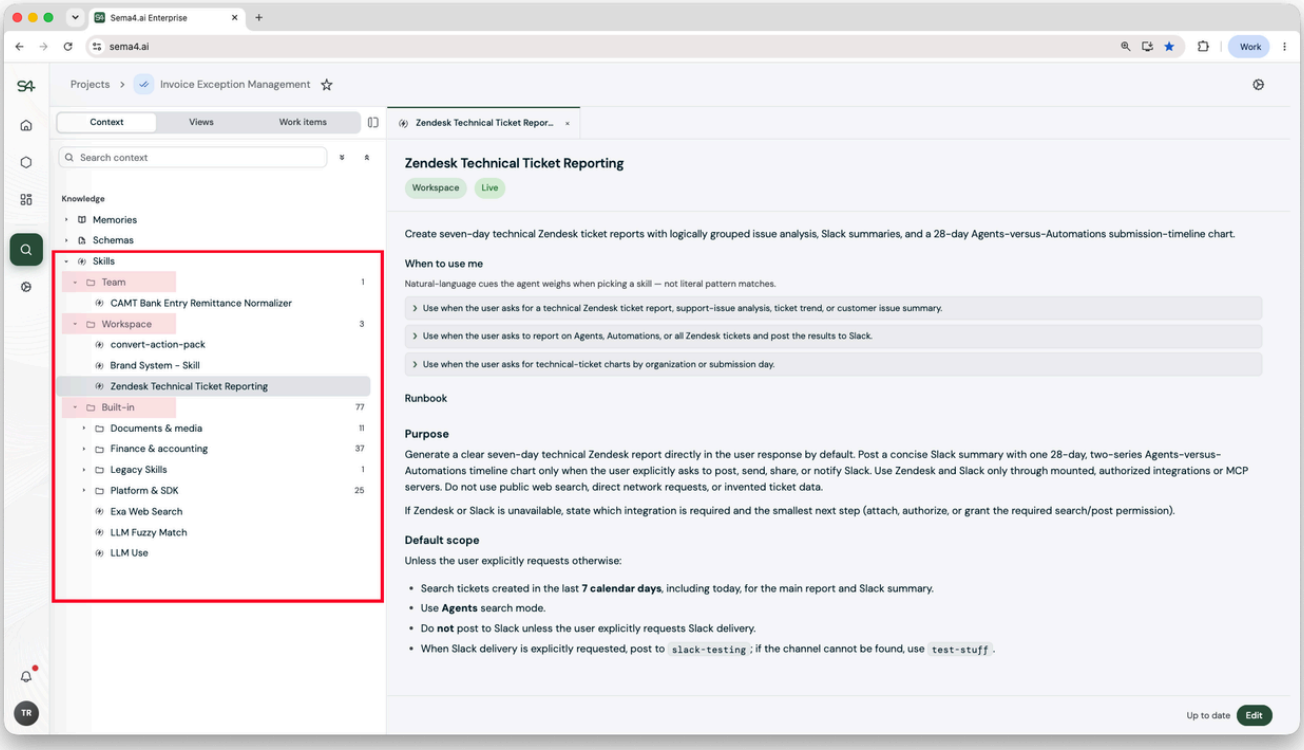
That is the focus of Sema4.ai v3.0. The platform helps enterprises move from isolated AI assistance toward team-owned AI work for the document- and data-heavy, multi-step processes that run the back office. It is designed to help organizations improve speed without weakening controls, scale operations without increasing manual effort at the same rate, and build AI workflows that business leaders can rely on.

Sema4.ai v3.0 adds stronger support for this reality with governed, shared workspaces for the assets teams use to build and run agents.

Shared workspaces for the work that teams own

Teams can share and manage workspaces, agents, skills, and files with control over how those assets are used. This helps move operational knowledge out of disconnected inboxes, spreadsheets, and one-off prompts into a more durable, collaborative operating environment.

Instead of each person maintaining their own version of a process, teams can work from shared assets that support more consistent execution and ongoing improvement.

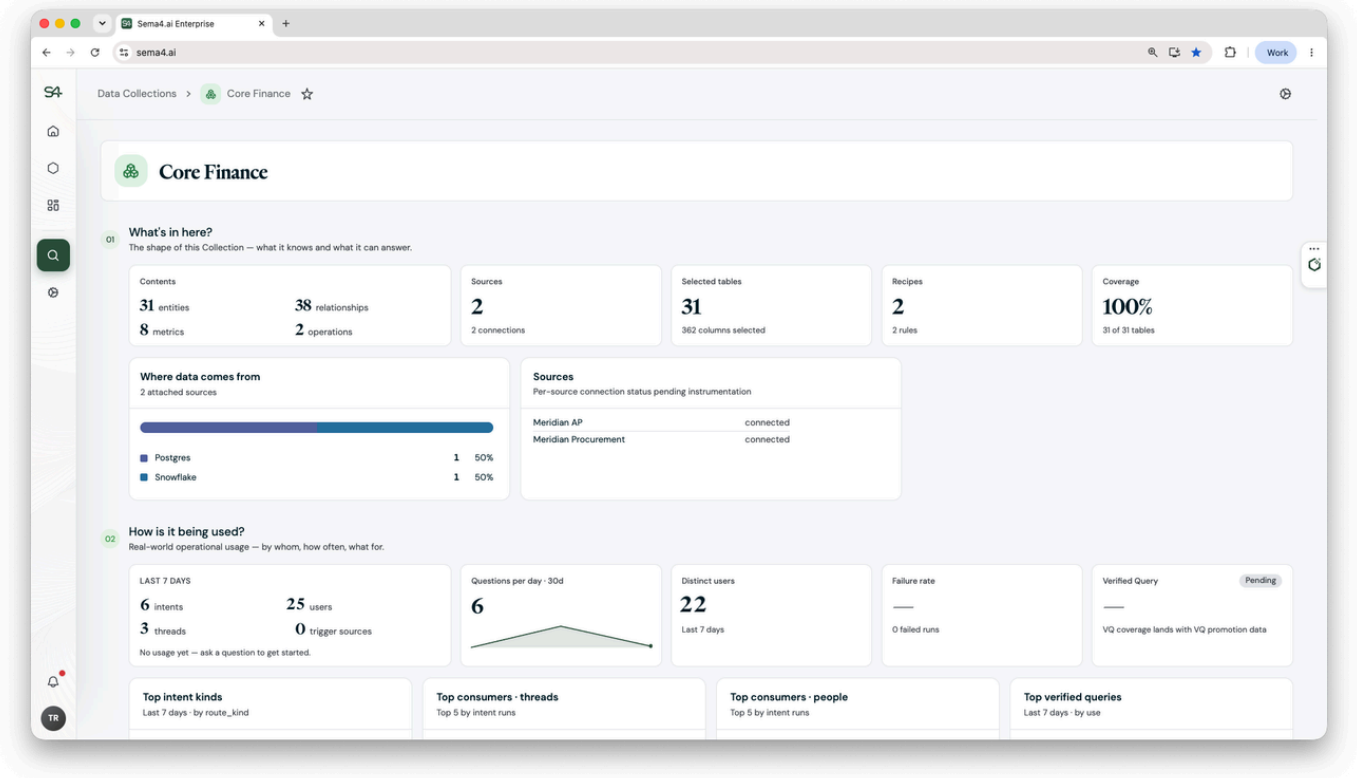

Sema4.ai Enterprise
sema4.ai
Work
Projects
Invoice Exception Management
Context
Views
Work items
Search context
Knowledge
Memories
Schemas
Skills
Team 1
CAMT Bank Entry Remittance Normalizer
Workspace 3
convert-action-pack
Brand System - Skill
Zendesk Technical Ticket Reporting
Built-in 77
Documents & media 11
Finance & accounting 37
Legacy Skills 1
Platform & SDK 25
Exa Web Search
LLM Fuzzy Match
LLM Use
Zendesk Technical Ticket Repor...
Zendesk Technical Ticket Reporting
Workspace
Live
Create seven-day technical Zendesk ticket reports with logically grouped issue analysis, Slack summaries, and a 28-day Agents-versus-Automations submission-timeline chart.
When to use me
Natural-language cues the agent weighs when picking a skill — not literal pattern matches.
Use when the user asks for a technical Zendesk ticket report, support-issue analysis, ticket trend, or customer issue summary.
Use when the user asks to report on Agents, Automations, or all Zendesk tickets and post the results to Slack.
Use when the user asks for technical-ticket charts by organization or submission day.
Runbook
Purpose
Generate a clear seven-day technical Zendesk report directly in the user response by default. Post a concise Slack summary with one 28-day, two-series Agents-versus-Automations timeline chart only when the user explicitly asks to post, send, share, or notify Slack. Use Zendesk and Slack only through mounted, authorized integrations or MCP servers. Do not use public web search, direct network requests, or invented ticket data.
If Zendesk or Slack is unavailable, state which integration is required and the smallest next step (attach, authorize, or grant the required search/post permission).
Default scope
Unless the user explicitly requests otherwise:
Search tickets created in the last 7 calendar days, including today, for the main report and Slack summary.
Use Agents search mode.
Do not post to Slack unless the user explicitly requests Slack delivery.
When Slack delivery is explicitly requested, post to slack-testing ; if the channel cannot be found, use test-stuff .
Up to date
Edit
TR

Shared business context with Data Collections

To be effective, agents need more than access to raw data; they need business context.

Our Data Collections capability brings related data sources together under a shared business view. A built-in Business Catalog helps agents work with terms such as suppliers, invoices, purchase orders, accounts, and their relationships, rather than leaving it to the LLM to infer meaning from disconnected database structures.

Teams can query data where it already lives while maintaining existing governance. They can also trace results back to the underlying source data.



Remember everything that matters to the business

Every business process has local knowledge: policy adjustments, customer requirements, exception rules, and operational nuances that are not always captured in a standard procedure.

Memories help teams preserve this context so agents can account for approved customer-specific guidance during future runs. That makes it easier to operationalize the knowledge that experienced teams apply every day.

Make repeat work more reliable and less expensive

AI models are probabilistic by design. They are valuable for interpretation, reasoning, and handling unstructured information. Yet repeatedly asking a model to generate the same data query, calculation, or transformation can introduce inconsistency and unnecessary cost.

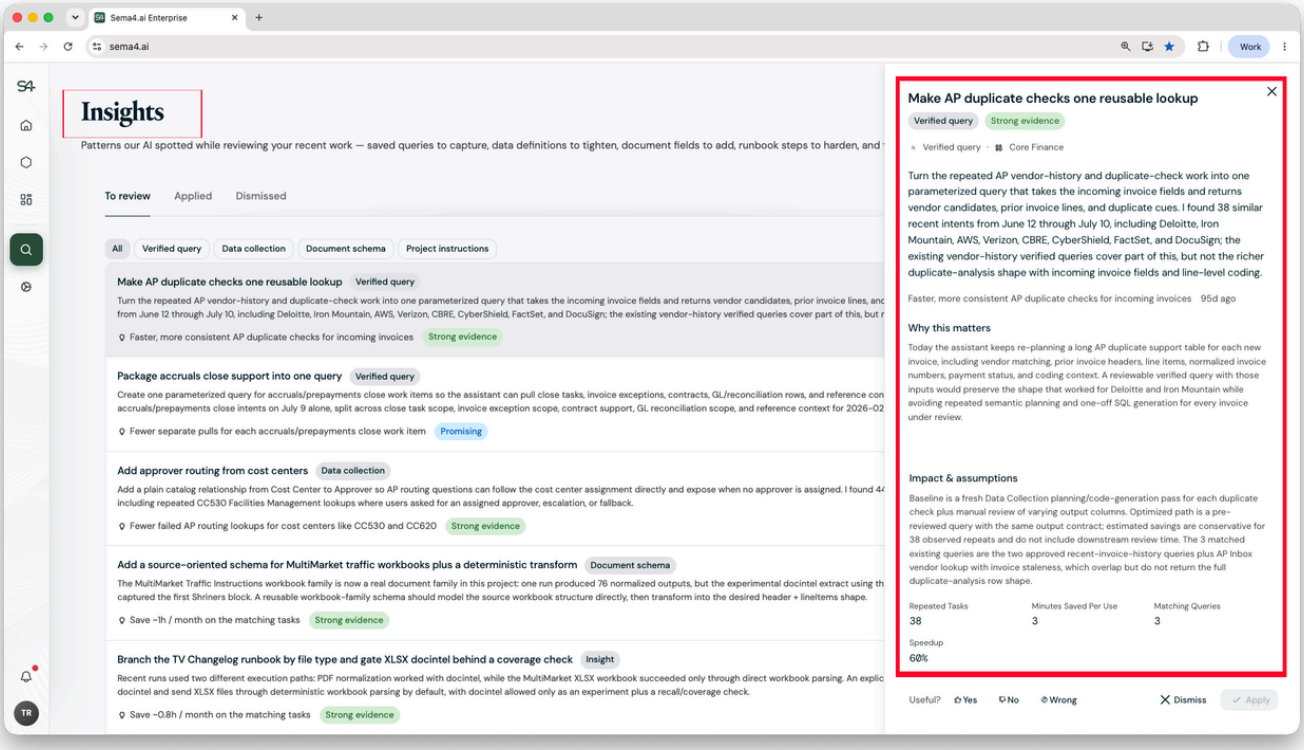
Sema4.ai v3.0 helps teams capture proven approaches as reviewed, reusable building blocks.

With our Verified Queries capability, teams can reuse approved data-access patterns for common questions. With our Verified Logic capability, they can reuse approved calculations, checks, and transformations. Each reusable artifact is reviewed, versioned, and tied to the work in which it runs.

This gives teams a more consistent foundation for repeatable operational work, including:

- Reconciliation and validation checks
- Data retrieval across approved business systems
- Rule-based calculations and transformations
- Exception handling that follows established policy
- Control-oriented workflows that require traceable execution

Rather than recreating the same approach from scratch every time, teams can retain and improve the logic that already works.


Sema4.ai Enterprise
sema4.ai
Work
Insights
Patterns our AI spotted while reviewing your recent work — saved queries to capture, data definitions to tighten, document fields to add, runbook steps to harden, and
To review
Applied
Dismissed
All
Verified query
Data collection
Document schema
Project instructions
Make AP duplicate checks one reusable lookup
Verified query
Turn the repeated AP vendor-history and duplicate-check work into one parameterized query that takes the incoming invoice fields and returns vendor candidates, prior invoice lines, anc
from June 12 through July 10, including Deloitte, Iron Mountain, AWS, Verizon, CBRE, CyberShield, FactSet, and DocuSign; the existing vendor-history verified queries cover part of this, but
Faster, more consistent AP duplicate checks for incoming invoices
Strong evidence
Package accruals close support into one query
Verified query
Create one parameterized query for accruals/prepayments close work items so the assistant can pull close tasks, invoice exceptions, contracts, GL/reconciliation rows, and reference con
accruals/prepayments close intents on July 9 alone, split across close task scope, invoice exception scope, contract support, GL reconciliation scope, and reference context for 2026-02
Fewer separate pulls for each accruals/prepayments close work item
Promising
Add approver routing from cost centers
Data collection
Add a plain catalog relationship from Cost Center to Approver so AP routing questions can follow the cost center assignment directly and expose when no approver is assigned. I found 44
including repeated CC530 Facilities Management lookups where users asked for an assigned approver, escalation, or fallback.
Fewer failed AP routing lookups for cost centers like CC530 and CC620
Strong evidence
Add a source-oriented schema for MultiMarket traffic workbooks plus a deterministic transform
Document schema
The MultiMarket Traffic Instructions workbook family is now a real document family in this project: one run produced 76 normalized outputs, but the experimental docintel extract using th
captured the first Shriners block. A reusable workbook-family schema should model the source workbook structure directly, then transform into the desired header + lineItems shape.
Save ~1h / month on the matching tasks
Strong evidence
Branch the TV Changelog runbook by file type and gate XLSX docintel behind a coverage check
Insight
Recent runs used two different execution paths: PDF normalization worked with docintel, while the MultiMarket XLSX workbook succeeded only through direct workbook parsing. An explic
docintel and send XLSX files through deterministic workbook parsing by default, with docintel allowed only as an experiment plus a recall/coverage check.
Save ~0.8h / month on the matching tasks
Strong evidence
TR
Make AP duplicate checks one reusable lookup
Verified query
Strong evidence
Verified query · Core Finance
Turn the repeated AP vendor-history and duplicate-check work into one parameterized query that takes the incoming invoice fields and returns vendor candidates, prior invoice lines, and duplicate cues. I found 38 similar recent intents from June 12 through July 10, including Deloitte, Iron Mountain, AWS, Verizon, CBRE, CyberShield, FactSet, and DocuSign; the existing vendor-history verified queries cover part of this, but not the richer duplicate-analysis shape with incoming invoice fields and line-level coding.
Faster, more consistent AP duplicate checks for incoming invoices 95d ago
Why this matters
Today the assistant keeps re-planning a long AP duplicate support table for each new invoice, including vendor matching, prior invoice headers, line items, normalized invoice numbers, payment status, and coding context. A reviewable verified query with those inputs would preserve the shape that worked for Deloitte and Iron Mountain while avoiding repeated semantic planning and one-off SQL generation for every invoice under review.
Impact & assumptions
Baseline is a fresh Data Collection planning/code-generation pass for each duplicate check plus manual review of varying output columns. Optimized path is a pre-reviewed query with the same output contract; estimated savings are conservative for 38 observed repeats and do not include downstream review time. The 3 matched existing queries are the two approved recent-invoice-history queries plus AP Inbox vendor lookup with invoice staleness, which overlap but do not return the full duplicate-analysis row shape.
Repeated Tasks
38
Minutes Saved Per Use
3
Matching Queries
3
Speedup
60%
Useful?
Yes
No
Wrong
Dismiss
Apply

Verified building blocks save time and money

Sema4.ai v3.0 improves cost efficiency by reducing redundant generation of work that can be completed using approved logic or structured data processing.

The platform's measured efficiency gains for repeat work include:

- **Verified Queries** reuse approved data-access patterns instead of generating a new query for each request: 80% fewer LLM tokens and 4× faster than free-form AI query generation.
- **Verified Logic** reuses approved calculations, checks, and transformations: 50% fewer LLM tokens and 30% faster than free-form AI-generated code.
- **DataFrames** process large, structured datasets without putting every row into a model context window, reducing unnecessary model use for data-intensive analysis.

Additionally, DataFrames give agents a structured workspace for calculations, comparisons, reconciliations, filtering, and transformations across large business datasets. This makes it more practical to complete a detailed analysis without treating the model context window as the primary place for data processing.

The end result is a more efficient path to scaling AI work, and the more you use it, the more you save. Sema4.ai uses models where judgment and interpretation help, and uses approved logic and structured processing where repeatability, speed, and cost matter most.

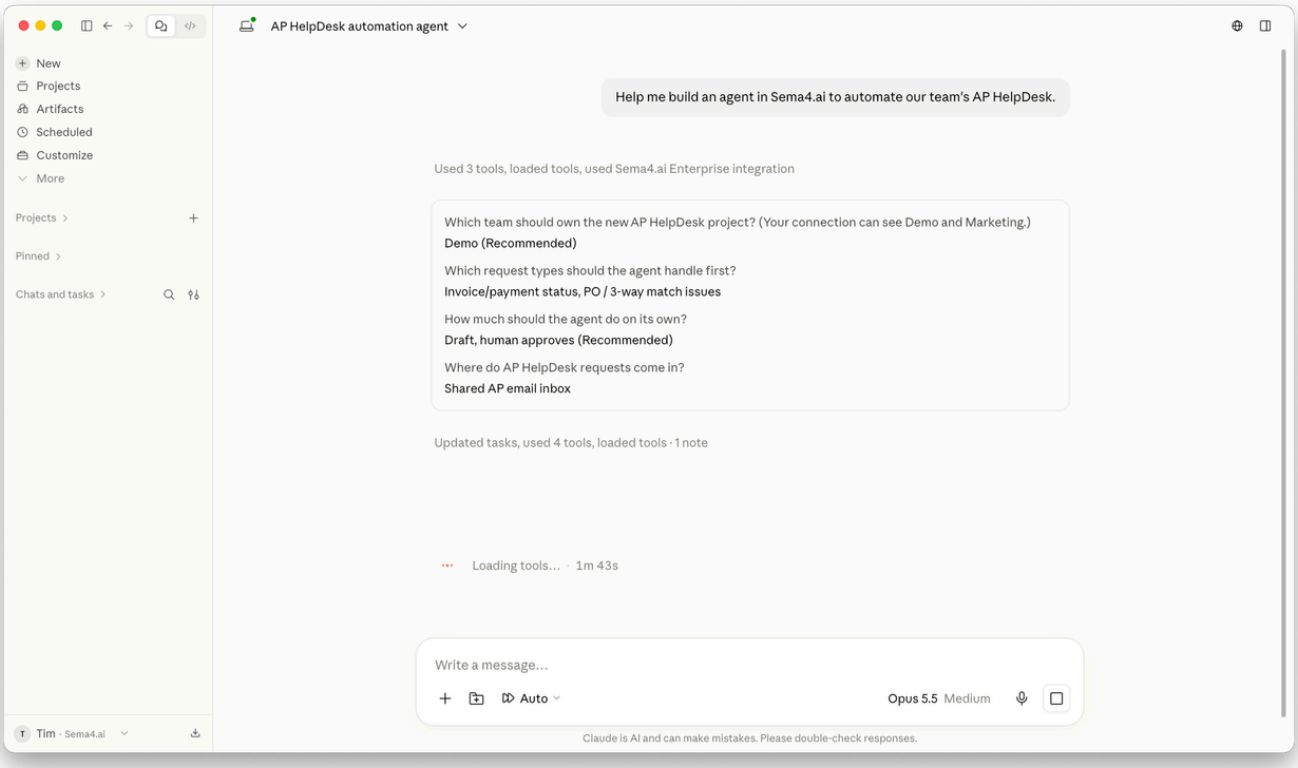
Integrate and easily work across the tools teams already use

Enterprise AI should fit into the environments where work already happens.

Sema4.ai v3.0 expands ecosystem interoperability by supporting AI personal productivity and business tools, including Claude Cowork, Microsoft Copilot, and Microsoft Fabric. Teams can create and engage with agents in their tool of choice while maintaining a shared foundation for governed operational processes.

The new Agent Inbox capability also helps teams monitor Microsoft 365 and Google Workspace email in one place, handle webhooks in real time, and route work using plain-language rules.

This creates a practical bridge between ad hoc requests and ongoing operational workflows. Teams can use familiar tools to engage with agents while bringing repeatable work into a governed, shared environment.



Support the full agent lifecycle

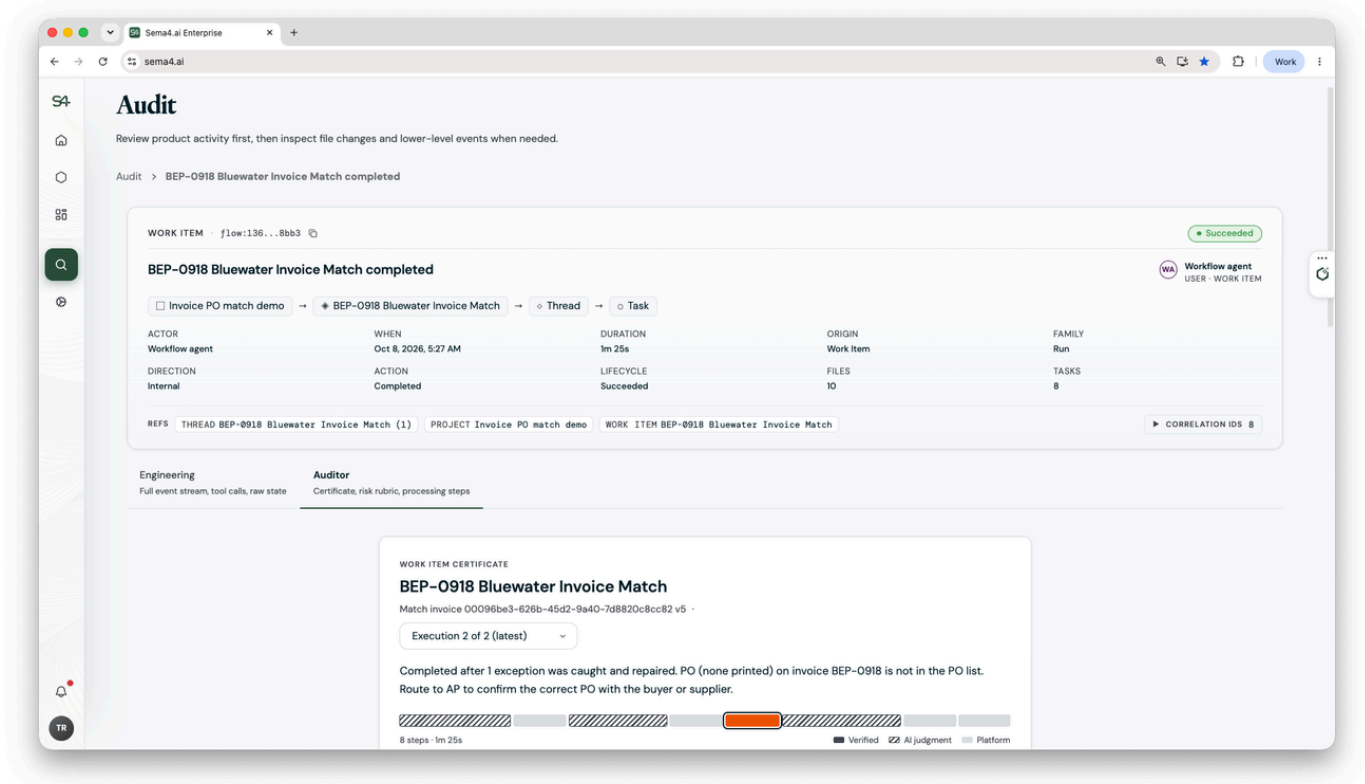
Building an agent demonstration is one challenge. Operating, changing, and auditing an agent over time and in real production use cases is another.

Sema4.ai v3.0 broadens support across the full lifecycle of enterprise AI agents:

- **Build:** Create and improve workflows, instructions, skills, document extraction, and data context.
- **Run:** Execute work across documents, data, business rules, and approved external tools.
- **Manage:** Review changes, investigate outcomes, and improve agents based on real usage.

Sai, the integrated AI assistant, is pervasive throughout the platform experience, helping teams use the dynamic, localized context of their current experience to improve workflows, instructions, skills, and data interactions.

Our enhanced Audit Logging provides a common record of work-item activity, documents, data, approved logic, and actions taken. Teams can investigate what happened, understand meaningful changes, and support SOX-ready control reviews with clearer evidence.



Overcoming the production gap for enterprise AI agents

Many organizations have made progress with AI assistants and early agent pilots. However, moving those pilots into ongoing business operations introduces a different set of requirements.

Teams need to address three practical challenges:

- **Trust and accuracy:** Repeated work needs dependable execution, clear controls, and evidence that teams can review.
- **Cost and efficiency:** Model use can grow quickly as workloads scale, especially when the same queries, calculations, and transformations are regenerated on every run.
- **Team-based governance:** Complex operational processes belong to teams, not individuals. The knowledge, rules, and assets behind that work need to be shared and managed accordingly.

Sema4.ai v3.0 addresses these needs by helping teams move more of their repeatable work from costly, probabilistic model generation into approved,

reusable capabilities.

Explore Sema4.ai v3.0

Sema4.ai v3.0 helps teams tackle complex, team-based work and dramatically reduce token usage costs by bringing together shared workspaces, reusable, approved logic, governed data context, structured data processing, interoperability, and audit-ready visibility into one platform for enterprise AI agents.

For teams looking to move beyond pilots, the next step is not simply adding more AI. It is creating a more reliable operating model for the work AI is asked to do.

Read the press release

Get a demo today.

Explore the Sema4.ai platform