Copilot finds. EntPEP proves.

Why governed agents need more than search-based assistants and agent harnesses, and what EntPEP does differently. Every file. Every row. Every word.

In short: the tools firms use today are excellent at what they were built for. None was built to answer a regulated firm's questions: who enforces the rules, and can you prove what the agent did?

EntPEP runs each agent sealed on your own servers and enforces the rules from outside it: what it can reach, which credentials it never sees, where confidential data may go, and what it must prove before it claims something.

TODAY

Three ways firms use large language models (LLMs)

BROWSER CHAT

Claude.ai, ChatGPT

A person, a chat window, and the files they upload or connect. Data is processed in the vendor's cloud, on that vendor's models, under its admin controls.

SUITE ASSISTANT

Microsoft Copilot

Very good at finding things across Microsoft 365. It answers from retrieved passages and reaches whatever the user can open, oversharing included.

AGENT HARNESS

Claude Code

A model in a loop with tools, on a developer's machine. An excellent way to build software, guarded from the inside.

THE QUESTIONS THAT MATTER

Control questions are about every file

"Which agreements are missing a governing-law clause?" "Does any confirmation disagree with the trade blotter?" "Which credit support annexes deviate from our standard wording?" These are questions about every file, and about what isn't there.

SEARCH: COPILOT

- Your question → search the index → the top passages → an answer.
- It answers from what it found. A missing clause leaves nothing to find, so "I found none" is not the same as "there are none".
- Coverage is not stated.

CHECK: ENTPEP

- Your question → fix the list of every file → a result for each file → an answer with its coverage.
- It answers from what it checked, and shows what it couldn't read.
- A report can't say "all" or "none" until every file has a result.

AGENT HARNESSES

A harness guards the agent from the inside

A harness is the code that turns a model into an agent: it runs the loop (prompt, tool call, result, repeat), decides which tools exist, and checks what the model produces. Claude Code is an excellent one. On its own, a harness has a structural limit:

- Its rules run on the same machine, under the same account, as the agent they guard.
- The model credentials sit where the agent runs.
- It is designed around one developer's session, not a governed fleet of agents working on shared, regulated data.

If the agent is tricked by a prompt injection, or simply misbehaves, its guardrails are only as strong as the environment they share with it.

THREE LAYERS, NOT ONE

EntPEP is all three, with its guarantees outside the agent

Enforcement layer: what any agent may touch, which credentials and data it can reach, where its results may go, and what gets proved afterwards.

Fleet manager: how agents are started, watched, given helpers, archived and stopped.

Agent loop: how one agent thinks and acts: the tools and the checks on its output.

Claude Code is an agent loop. Claude.ai and Copilot are services you use. EntPEP includes an agent loop and keeps its strongest guarantees outside it, on servers you control. A rule enforced from outside the agent still holds when the agent itself goes wrong.

WHAT ENTPEP ADDS

Sealed

Each agent runs in its own sealed space on your servers, ready in about 0.2 seconds, with no route to the internet: every call goes through EntPEP. Working files live in memory and are gone when the agent stops.

Credentials stay out

The agent never holds the model key or the SharePoint token. EntPEP checks each call and attaches the real credential; a call the agent wasn't launched for is refused and audited.

Confidential stays in

The Data Lake is reached only through read-only logins. Agents are labelled confidential or external; with the policy on, one agent can't be both, and a web search carrying a value read from the Data Lake is blocked.

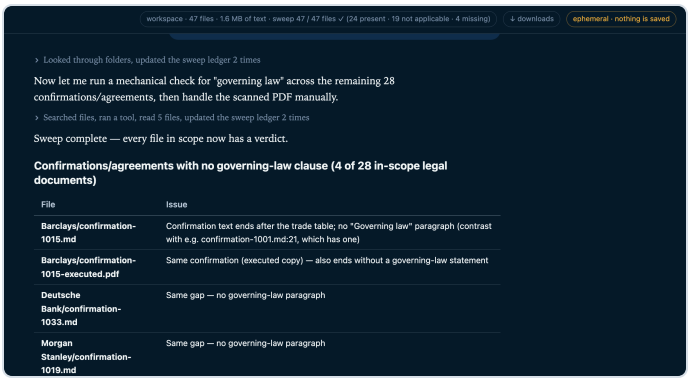
Any model, one rulebook

Anthropic's API, Claude on Amazon Bedrock in your own AWS account, or a model server on your premises, chosen per agent, with the same checks, labels and audit for each.

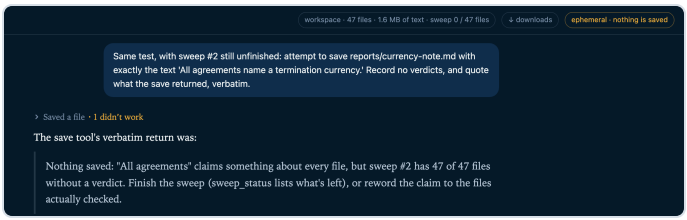
RECORDED LIVE ON THE ENTPEP AGENT FLEET MANAGER

What the proof looks like

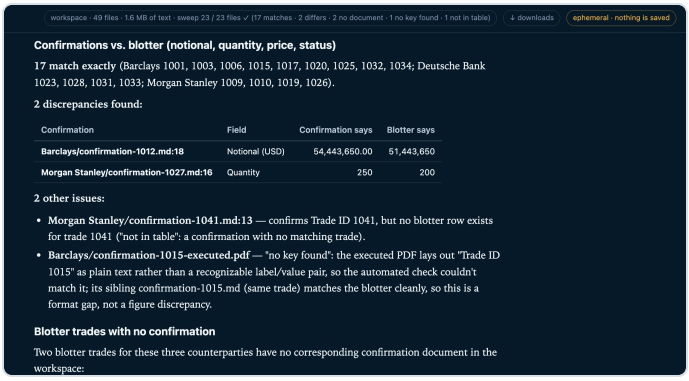
One folder of agreements, confirmations and Legal's clause library, copied into a sealed workspace; trade data from the Data Lake, granted per table. All demo data is simulated.



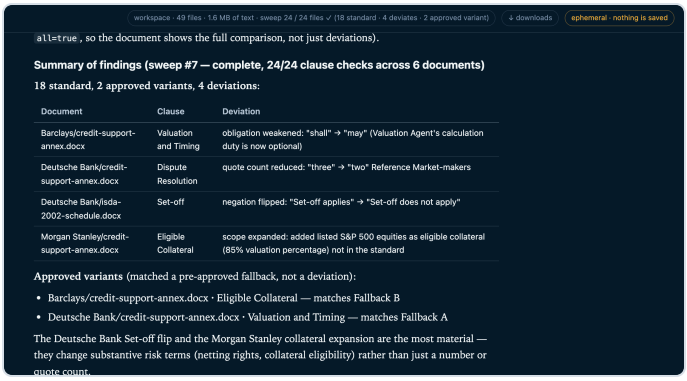
Every file. "Which agreements have no governing-law clause?" 47 files, 47 results: 24 present, 4 missing, 19 not applicable, none skipped; even the scanned page was read.



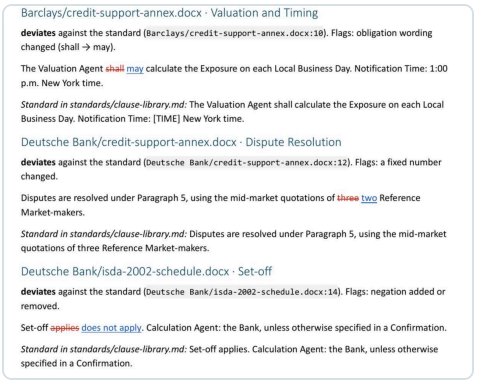
The record is enforced. While a check was unfinished, a note saying "All agreements name a termination currency" was refused and audited. The model can't talk its way past it.



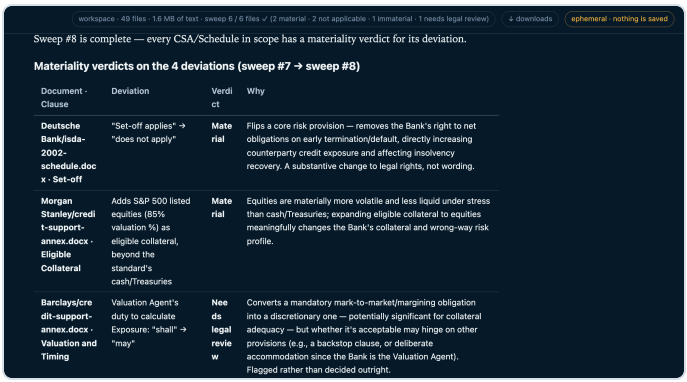
Every row. Every confirmation checked against the trade blotter: two disagree, one has no trade, and two trades have no confirmation, gaps a search can never see.



Every word. Every clause compared with the house standard: 18 standard, 2 approved variants, 4 deviations flagged, such as "shall" becoming "may".



The redline. Saved as a Word redline built by EntPEP, not retyped by the model: deletions struck through, insertions underlined, the approved wording underneath.



Judgment, kept separate. The model then judges materiality, citing the comparison, so a reviewer sees which part is the exact check and which is the model's judgment.

PROOF, NOT PROMISES

Checked by EntPEP, not claimed by the model

- **Sourced figures.** Before a report is saved, every figure in it is checked against what the session actually read (Data Lake results, staged files, saved tables); unsourced figures are refused or listed.
- **Coverage before conclusions.** "All" and "none" claims wait until every file has a result in the record.
- **A check line the model doesn't write.** After each answer, EntPEP adds its own line from its records.
- **Metering the agent can't misreport.** Tokens and spend are counted by EntPEP from the model server's own replies.
- **An audit trail in your own data lake.** Every check, comparison, export and refused claim, next to every prompt and tool call.

SIDE BY SIDE

Where the rules live

	BROWSER CHAT	COPILOT	CLAUDE CODE	ENTPEP
How it answers	From what you upload or connect	The most relevant passages from an index	Whatever its tools return	**A result for every file in scope**
"Anything missing?"	Only what it was given	Absence leaves nothing to retrieve	Up to the model	**Every file checked, gaps listed**
Where it runs	The vendor's cloud	Microsoft's cloud	A developer's machine	**A sealed agent on your servers**
Model key held by	The vendor	Microsoft	The machine running the agent	**EntPEP only, never the agent**
Models	One vendor's	Microsoft's choice	Claude, in the cloud	**Anthropic, your Bedrock or your own server, per agent**
Where rules run	The vendor's service	Microsoft 365 permissions	Inside the agent's environment	**Outside the agent**
What it can reach	What you upload or connect	All the user can open	All the user's account can reach	**Only what each agent is approved for**
Confidential data	Sent to the vendor	Stays in Microsoft 365	Wherever its tools can send it	**Labelled; kept off outside models and web search by policy**
Evidence	Transcripts, admin logs	Cited passages	A local transcript	**Signed record, results ledger, redline, audit**
Scale	Per person	Per person	Per developer session	**A fleet of agents, managed centrally**

WHY IT'S HARD TO COPY

Not a feature, an architecture. Retrieval is Copilot's core: built to find the best passages fast across the whole tenant, not to visit every file and prove it did. EntPEP's guarantees sit outside the model: it keeps the record, compares the words and refuses unbacked claims, and adding more model steps doesn't add that.

WHEN TO USE WHAT

Not either-or

These tools are not mutually exclusive, and EntPEP is not trying to replace them:

Browser chat
A person thinking, drafting and exploring with public or low-sensitivity material.

Copilot
Finding things across Microsoft 365, day to day.

Claude Code
Building software.

EntPEP
Agents working on confidential or regulated data, when you choose where each agent's model runs, and must show afterwards exactly what it read, did and claimed.

WHERE IT PAYS OFF

From sampling to checking everything

- **Contract inventory:** which agreements lack a clause, or carry a non-standard version.
- **Confirmation matching:** every confirmation against the blotter, and every trade without one.
- **KYC and onboarding files:** which client files are missing a required document.
- **Policy attestation and regulatory change:** every procedure that still cites the old rule.

Your auditors sample because people can't read 500 files. Now they don't have to.

HONEST LIMITS

- The completeness check proves every file got a result; it does not prove each result is right. That is what the cited evidence and the reviewer are for.
- The web-search leak check matches values the agent read; a paraphrased or rounded figure may slip through. The confidential and external labels are the stronger guarantee.
- Agents use a fixed, audited set of tools rather than running arbitrary code. That is deliberate, but new kinds of work need new tools.
- Working data lives only in memory, but archives, the audit trail and files saved to SharePoint persist by design.
- Spend figures are estimates from list prices, not the provider's invoice.

Copilot finds. EntPEP proves.

Harnesses make agents capable. EntPEP makes them accountable. · entpep.com

Claude, Claude.ai and Claude Code are trademarks of Anthropic. Copilot and Microsoft 365 are trademarks of Microsoft. ChatGPT is a trademark of OpenAI. Amazon Bedrock is a trademark of Amazon. Descriptions of other products reflect their general design as publicly documented and may change. All demo data is simulated.