

Building an AI-Native Data Platform

Data platforms gave every knowledge worker governed data. Its meaning lives in people’s heads, out of agents’ sight. Generative AI now enables the next generation: a knowledge layer for both.

The knowledge staircase

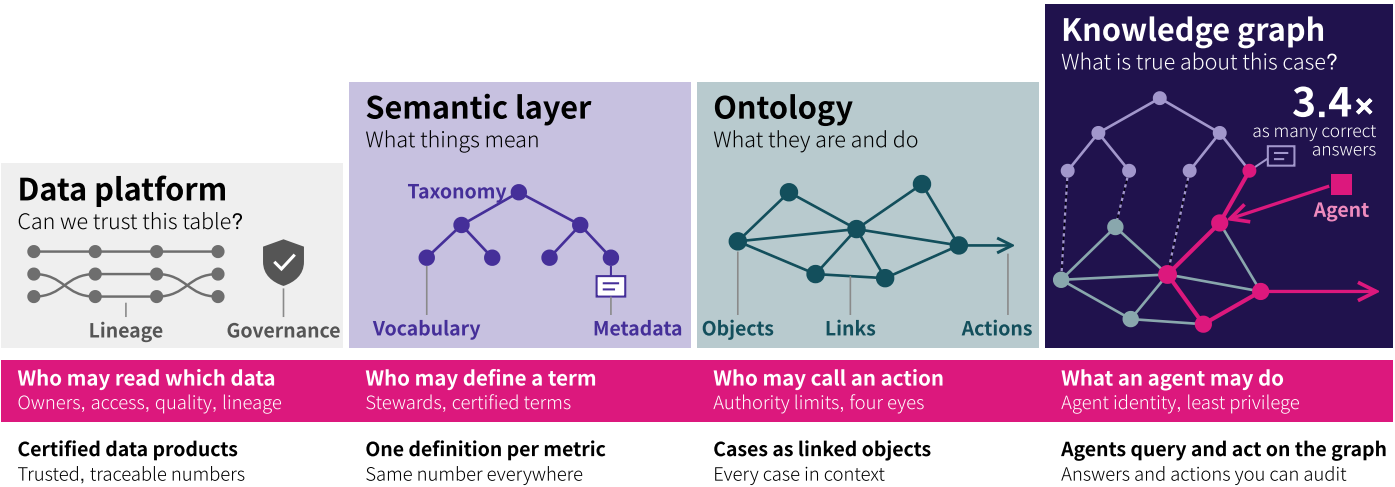


Fig. 1 An agent starts at a case, reads its terms, follows the links and leaves with an action. Each step adds its own governance. 3.4×: Sequeda et al., 2023.

From data to knowledge

Agents now read the same tables as the claims handlers, underwriters, actuaries and controllers they work beside. Those tables hold governed, trusted data. What still separates humans from agents is real-world context: a claims handler knows what a figure means, which table holds it, how a case links to its contract. Organisations have tried to close that gap for many years (Figure 2) through active alignment by central teams. Language and definitions keep changing, so those teams were always fighting an uphill battle to stay current. Generative AI changes that: always-on agents enrich the semantic layer and the ontology from real interactions, and the central team reviews. That puts an old dream within reach: a CEO asks for the cost per claim on a Friday evening and can get the answer independently without pulling someone else in to get the details. For the company, that means lower cost per case, faster decisions and auditable automation.

The semantic layer

The semantic layer says what things mean. It levels the playing field, putting the subject-matter expert and every agent into the same context: for example, how last year’s profit figures were calculated or what lies behind the KPIs. Jessica Talisman’s ontology pipeline defines the stepping stones, from controlled vocabulary to thesaurus (Figure 1). Generative AI now pulls these steps together, keeping information up to date for both humans and agents.

The ontology

The ontology adds the business objects and their relations, so agents see invoice, contract and policyholder as one case, as a claims handler does. Its links serve three jobs. Act: check the cover and approve an invoice within the agent’s limit. Detect: flag a provider who bills the same treatment twice. Trace: follow a contract change forward to every open claim it affects, or a rising cost per claim back to the providers behind it.

The knowledge graph

Imagine your city’s public transport: lines connect the stops, and around each stop lies a neighbourhood buzzing with context. Put the semantic layer and the ontology together and you get the same: the knowledge graph. Business objects such as contract and invoice are its stops, their relationships its lines and the definitions its neighbourhoods. An agent starts at a case, reads what its terms mean, follows the links and leaves with an action (Figure 1). In a benchmark by Sequeda, Allemang and Jacob (2023), a language model answered 54% of enterprise questions correctly with a curated and verified knowledge graph, against 16% with the raw tables alone.

In the knowledge graph, the semantic layer says what things mean and the ontology says how they connect, all on a trusted data platform.

Two traditions, one knowledge layer

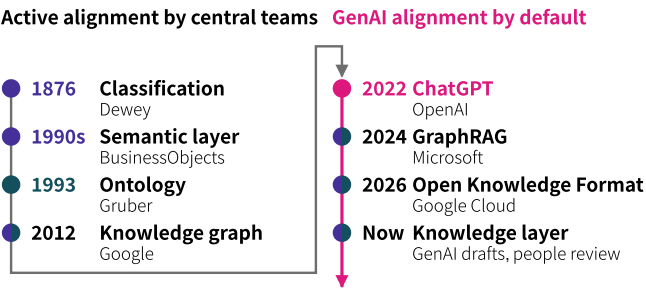


Fig. 2 On the left, every step needed a central authority to keep it aligned: librarians, BI teams, Google’s curators. On the right, language models draft terms and links. Violet: semantic. Teal: ontology. Split: both.

Both layers read the same serving data and meet in one graph. The ontology enables agents to perform governed write-back onto the data platform.

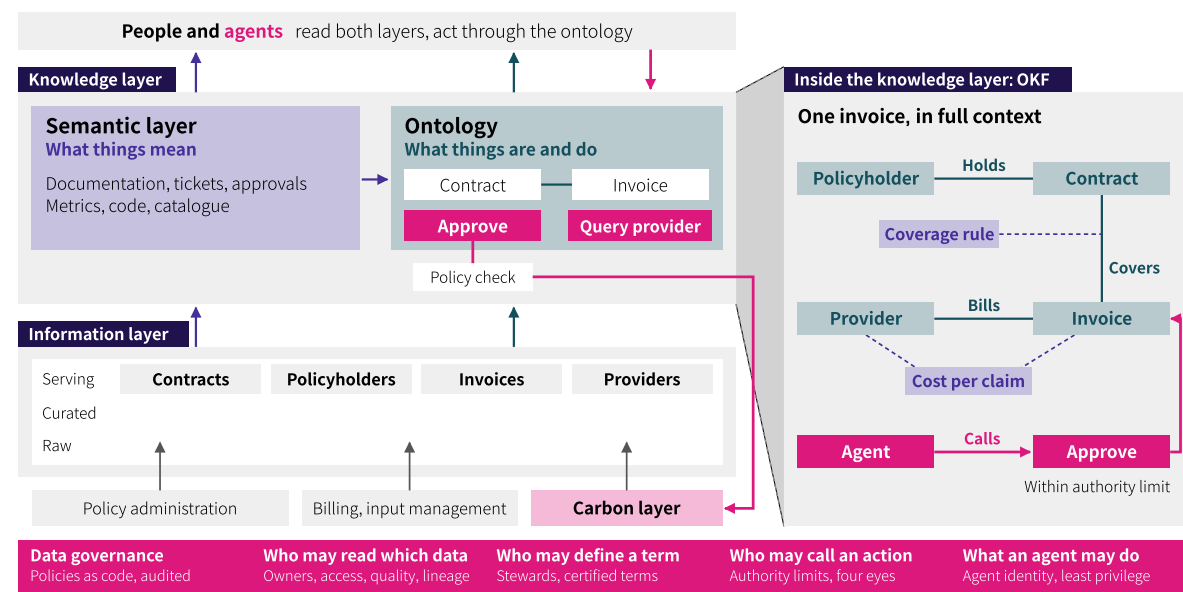


Fig. 3 Follow one invoice from billing to the Invoices data product. The ontology links it to its contract and provider, and the semantic layer adds the coverage rule and cost per claim. An agent approves it within the authority limit, and the approval returns via the policy check and carbon layer. Data governance runs under every step.

Built on the platform you run

The knowledge layer builds on the platform the company already runs. Raw, curated and serving data stay where they are, and both layers read the serving data and meet in one graph (Figure 3). Strict governance allows the ontology controlled write-backs through policy gates into the carbon layer, which the platform ingests like any other source.

How knowledge flows

The knowledge layer stays up to date without a central team, because knowledge enters where people already write it, from documentation and tickets to code and the catalogue, and through the data pipelines. Each data product arrives with drafted definitions. Generative AI drafts the definitions in version control, and a person reviews each one before it enters the graph. Every definition traces back to its source, author and reviewer (Figure 4). This closes the loop from incoming knowledge to governed action by an agent and a recorded outcome.

Knowledge moves in a loop

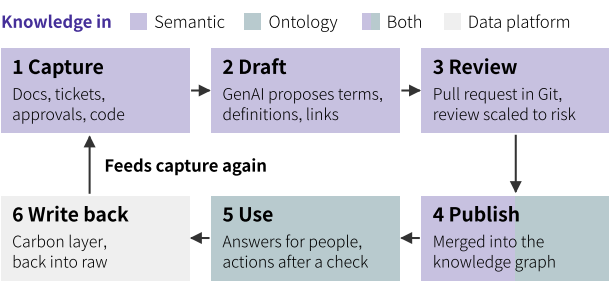


Fig. 4 Git records every step of the loop.

Where to start

Start from the value: reduced development times, faster insights for subject-matter experts or autonomous, AI-native processes are just three of many examples. Pick one: this will be the lighthouse to guide you.

- 1 The data platform:** build governed data products and curate them in a data catalogue, with owners, descriptions, lineage and data quality checks. If your curated layer is a well-designed Data Vault, its hubs and links already hold your business objects and their relationships.
- 2 The semantic layer:** have a look at your data contracts, your code or your table descriptions. Select the set your value case needs and bring it into a semantic graph (see Google’s Open Knowledge Format and Andrej Karpathy’s LLM Wiki).
- 3 The ontology:** take an industry ontology to accelerate the start. If your curated layer is a well-designed Data Vault, use it: its hubs and links already hold your business objects and their relationships. Your experts, with GenAI, will close the remaining gap to the organisation’s own objects and links.
- 4 The knowledge graph:** activate the knowledge. Bring all the parts together and open it to your people and agents for your value case. Trace their feedback and every outcome back into the graph, so the layer learns with each case.

What is the value for you?

Philipp Warmer Data and AI Lead
 Builds data platforms, GenAI solutions and governance for agentic AI in regulated industries. Dr. sc., ETH Zurich.
philipp.warmer@d-one.ai

Grant Kalil Managing Director Germany
 Over 15 years in go-to-market and enterprise transformation, most recently generative and agentic AI. MBA, HSG.
grant.kalil@d-one.ai

Sources J. Talisman, The Ontology Pipeline, Modern Data 101, 2025 and 2026. J. Sequeda, D. Allemang, B. Jacob, knowledge graph benchmark on enterprise SQL, arXiv, 2023. T. Gruber, 1993. Google Cloud, Open Knowledge Format, 2026. A. Karpathy, LLM Wiki, GitHub, 2026.