

Semantic trust transforms interoperability into AI advantage

Poor data undermines healthcare AI ROI, but a semantic trust flywheel turns it into a compounding advantage with the right architectural investment.



Finding ‘better’ AI models isn’t the challenge; data normalization and signal integrity are the levers that determine your interoperability ROI, and semantic trust is the one architectural commitment that leads to compounding AI outcomes.

Healthcare built interoperability infrastructure for 10 years. What changed?

- The investment was necessary. Interoperability worked. Data now moves between providers, pharmacies, payers, and labs at an enormous scale.
- Yet, the AI programs sitting on top of all that connectivity keep disappointing.
- Models underperform. Predictions go wrong in inexplicable ways. Clinicians stop trusting the outputs, and strategic growth initiatives stall.
- The diagnosis most executive teams reach is a model problem. The actual diagnosis, almost universally, is a data problem.
- Specifically, a semantic trust problem. ‘Connected’ doesn’t mean consistent. Building AI with inconsistent data is a massive strategic risk.

The core problem: Connected doesn’t mean consistent

Consider what ‘interoperable’ means in most enterprises today. A medication record for a single patient arrives from the EHR as ‘Metformin 500mg oral tablet’. The pharmacy system calls it Metformin HCl 500. The claims file records Metformin Tab 500MG. A fourth system carries the NDC code with no text description at all. These are four representations of identical clinical reality. In the absence of a normalization layer, an AI model trained on this data learns to recognize the encoding quirks of each source system, not the underlying clinical truth. That is not an edge case; it is the standard condition of healthcare data at rest. Fewer than four in ten health information exchanges routinely send data aligned with USCDI standards. Fewer than one in three routinely receive it. The connectivity problem has been substantially solved. The consistency problem has barely been touched.

Four ways this breaks AI

When AI systems are built on semantically inconsistent data, the failures follow a predictable pattern.

- **Signal loss:** Occurs because clinically significant events are missed when the same condition is coded differently across source systems, and no reconciliation layer catches the variants before training.
- **Duplicate identities:** This means a patient appears as multiple distinct records across provider, payer, and pharmacy systems, so the model treats one person's longitudinal history as several unrelated cases.
- **Provenance blindness:** Ensures the model cannot explain its predictions because there is no audit trail connecting outputs back to source transactions.
- **Ontological drift:** Standard terminologies evolve, and a model trained on last year's mappings silently degrades as reality moves away from it.

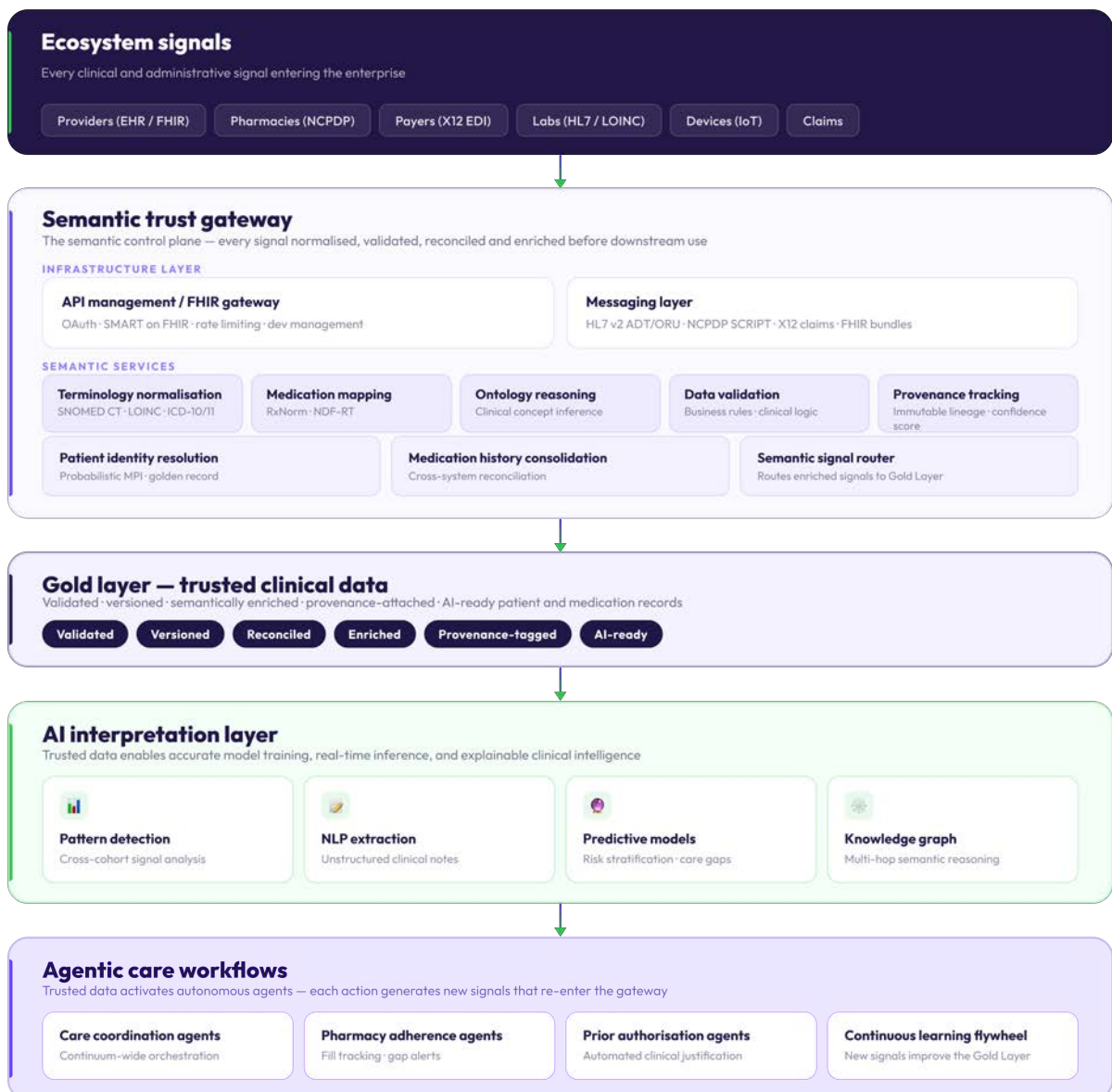
None of these are model failures. They are data infrastructure failures.
Addressing them requires architectural investment, not model retraining.

Semantic trust is not a general aspiration about data quality. It is a precise property of a data element. One that is normalized against a current standard terminology, validated against clinical logic, reconciled across source systems, and annotated with immutable provenance. When all four conditions are met, the element is fit for AI. When one fails, it is a liability. The response to this problem is not a product. It is an architectural posture. A deliberate sequencing of capability that ensures every clinical signal entering the enterprise is semantically trusted before it reaches any analytics, model, or automated workflow. We call this architecture the Semantic Trust Flywheel.

The semantic trust flywheel

The flywheel has five layers, and the sequence is not optional. Each layer is a prerequisite for the one above it.

Signal capture → semantic trust → AI acceleration: a self-reinforcing cycle



Layer 1: Capture every signal from every source

It requires comprehensive ingestion—every clinical and administrative signal from every relevant source. Provider EHR systems, pharmacies, payers, labs, connected devices, and remote monitoring streams must be captured. The semantic trust problem begins the moment these signals arrive, each carrying its own local encoding conventions and patient identifier scheme.

Layer 2: Semantic trust gateway—where meaning is established

The heart of the architecture—it sits between raw signal ingestion and every downstream consumer. It does four things, in sequence, at the point of ingestion.

- **Terminology normalization:** Maps every clinical concept to its canonical representation in the appropriate standard ontology. SNOMED CT or ICD-10/11 for diagnoses, LOINC for observations, RxNorm for medications with dose-form and strength disambiguation. This normalization runs against live terminology repositories, not static snapshots. When SNOMED CT updates, the gateway updates with it.
- **Ontology reasoning:** Enriches the normalized record. A patient coded with a SNOMED CT concept for Type 2 Diabetes Mellitus is automatically inferred to have membership in the broader diabetes mellitus class, enabling population queries and care gap identification that would otherwise require bespoke coding for every query.
- **Reconciliation:** Produces a single golden record where the same clinical concept arrives from multiple source systems. The gateway uses probabilistic and deterministic master patient index matching across all sources to resolve duplicate identities and consolidates medication histories across pharmacy, EHR, and claims sources. There is one patient. There is one medication history. That is what downstream systems see.
- **Provenance tracking:** Annotates every normalized data element with immutable lineage metadata. Which source system it came from, when it arrived, which terminology version was used to map it, the confidence score on that mapping, and any reconciliation that was applied. This chain is the foundation of AI explainability. When a predictive model fires a care gap alert, the clinical team can trace every contributing data element back to its source transaction.

Layer 3: The gold layer (the only substrate AI is permitted to train on)

It is the curated, versioned, continuously validated clinical data store that sits downstream of the gateway. It is the single authoritative representation of every patient, medication, encounter, observation, and claims event in the enterprise—normalized, reconciled, and provenance-tagged. Automated data quality checks run on every batch. No element enters without passing the four semantic trust conditions. This is the only data that any model in the enterprise trains on or infers from. The gold layer is meant to serve as a guarantee. That why what AI reasons over is worth reasoning over.

Layer 4: A trustworthy foundation. Accurate AI

With a semantically trusted gold layer beneath it, AI delivers what it has been promising. Pattern detection across normalized cohorts identifies comorbidity clusters and readmission precursors without terminological noise. NLP models extracting concepts from clinical notes are grounded against the same standard terminologies that govern structured data, so structured and unstructured signals can be analyzed together. Predictive models for deterioration, medication non-adherence, and care gap likelihood train on data whose integrity is assured. And a semantic knowledge graph—connecting patients, conditions, medications, providers, encounters, observations, pharmacy fills, and claims events through typed relationships—enables the kind of multi-hop reasoning that is simply impossible in a tabular data architecture.

That last capability deserves a concrete illustration. Consider the query: identify patients with a confirmed type 2 diabetes diagnosis who have been prescribed Metformin but have had no pharmacy fill in 90 days, and whose most recent HbA1c exceeds 8.0%. In a tabular model, this requires joining multiple tables across systems using identifiers that may be inconsistent. In a semantic knowledge graph built on trusted data, it is a four-hop traversal filtered by date and value at each node. The answer is available in real time. The clinical team can act on it today.

Layer 5: Agents that act and a flywheel that never stops

The apex of the architecture is agentic care: autonomous workflows that execute on behalf of clinical teams. Care coordination agents orchestrate transitions across the care continuum. Pharmacy adherence agents monitor dispensing records and alert on fill gaps correlated with deterioration signals. Prior authorization agents assemble clinical justification packages from the gold layer and submit them to payers without manual intervention.

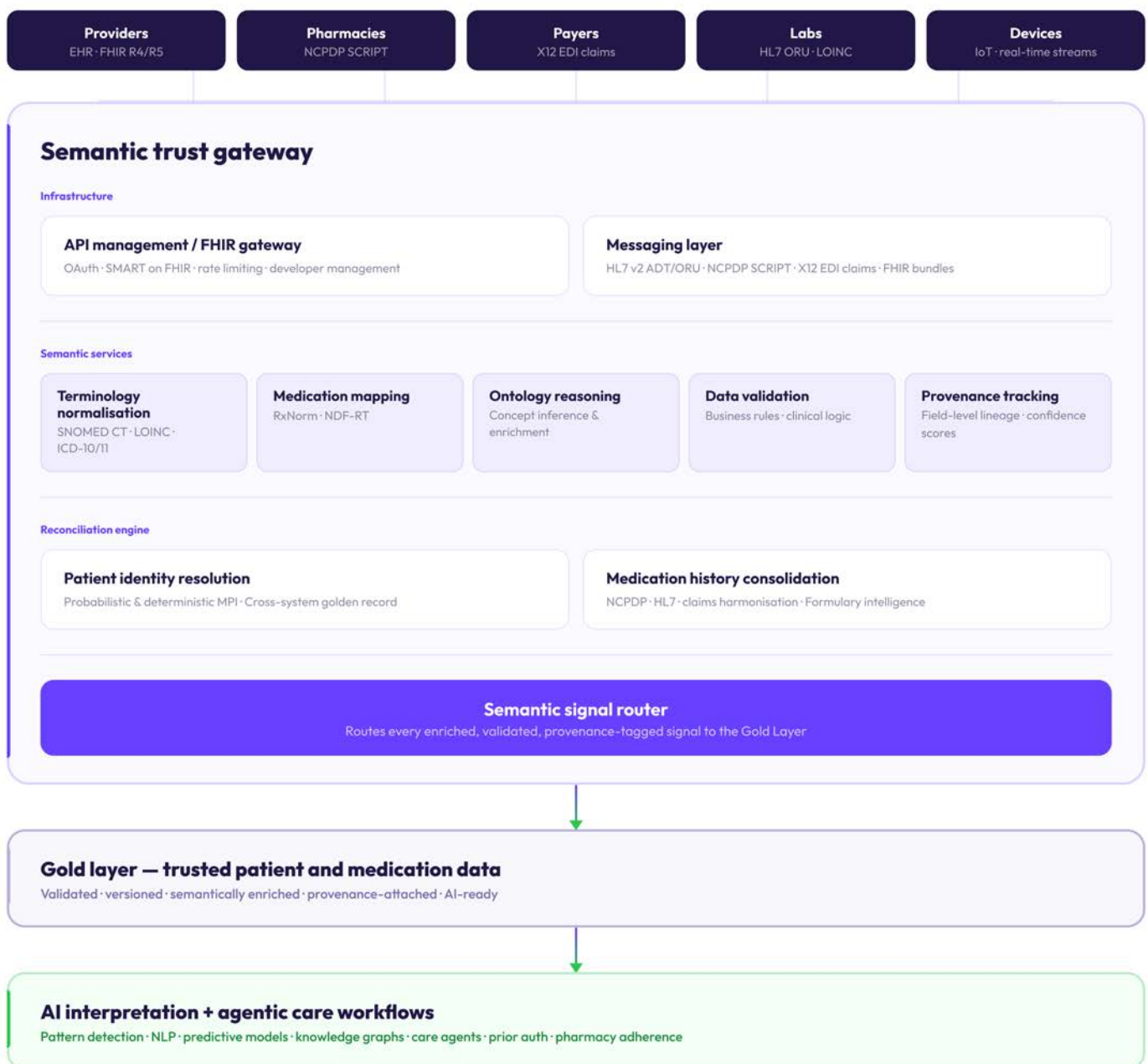
Each of these agent actions generates new clinical signals: an auth approval, a care gap closure, a pharmacist intervention record. Those signals flow back into the semantic trust gateway, are normalized and provenance-tagged, and enter the gold layer enriched. The flywheel does not stop. Every cycle improves the data foundation for the next one.

The maturity map: Where you are and where you need to be

Most large health systems and payers are operating at Stage 2 of a five-stage maturity model. They have working API connectivity, FHIR servers, HL7 ingestion, and data lake pipelines. The strategic investment decision in front of them is Stage 3—the semantic trust gateway—which is the prerequisite for every AI capability above it. This does not mean AI work must stop until the entire enterprise data estate is semantically clean. It means AI initiatives should be scoped to the clinical domains and patient populations for which semantic trust has been established. Stage 3 is not a phase; it is the foundation.

Unified data gateway — semantic architecture

Full component map: from raw ecosystem signals to AI and agentic care workflows



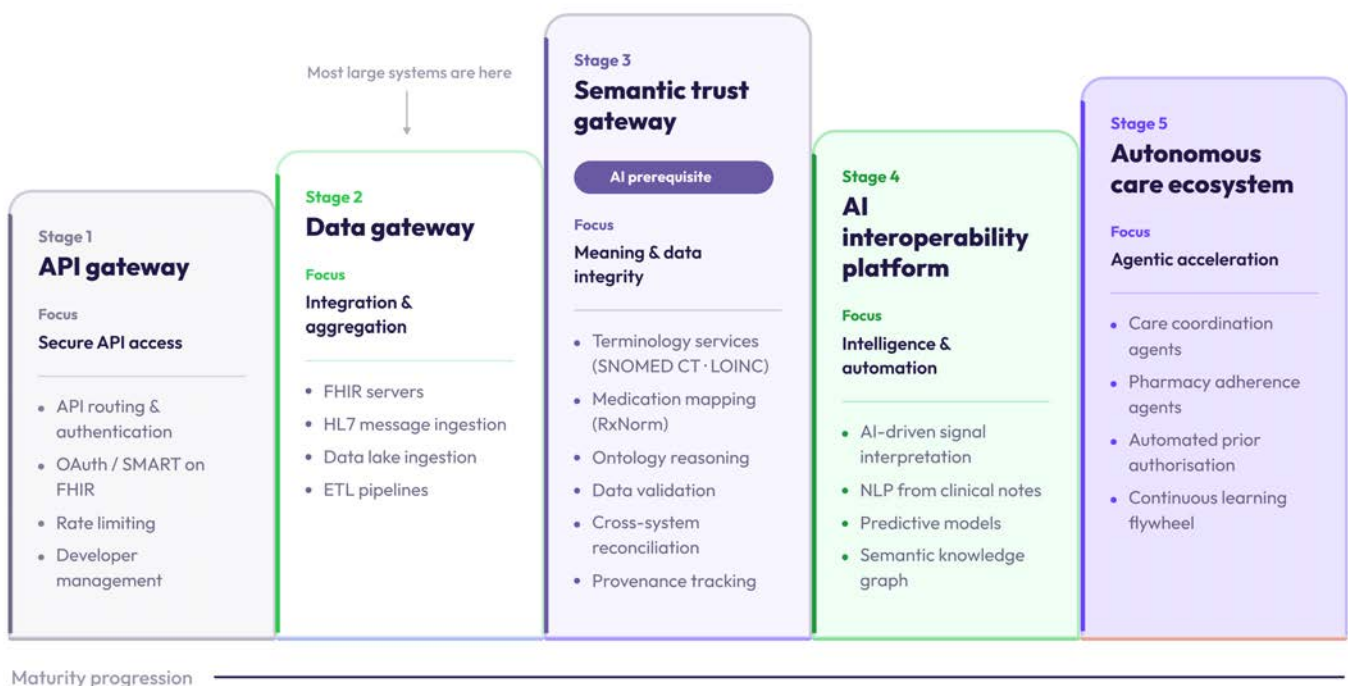
Semantic services execute at ingestion — every downstream consumer receives trusted data, not raw signals

The knowledge graph: A reasoning layer

The knowledge graph closes the loop. It sits above the gold layer—the reasoning substrate that makes the AI interpretation layer possible. It represents the full clinical ecosystem as a network of typed, provenance-tagged entities connected by semantically meaningful relationships. Patients have conditions. Conditions are treated by medications. Medications are dispensed by pharmacy fills. Encounters generate observations. Observations trigger care pathways.

Interoperability and AI maturity model

Five stages — Stage 3 is the prerequisite for every AI initiative above it



Without Stage 3, every AI initiative above it is building on semantically inconsistent data — technical debt accumulates faster than value

This graph is also the mechanism by which AI predictions become explainable. When a predictive model identifies a patient as high-risk for 30-day re-admission, the explanation is a subgraph—the specific constellation of conditions, medications, observations, and encounter patterns the model weighted most heavily. Because every node in that subgraph is provenance-tagged to its source transaction, the clinical team can review the underlying evidence. The prediction is not a black box. It is a navigable structure with a documented audit trail.

That explainability is not a nice-to-have. As regulatory frameworks governing AI in clinical decision support crystallize, the FDA's evolving position on AI/ML-based software-as-a-medical device is the most consequential. The ability to trace every AI-generated recommendation back to its source data will become a compliance requirement. Organizations that build provenance into their architecture now are investing in a regulatory asset. Those that defer it will face remediation costs and timelines far exceeding the cost of building it correctly from the start.

Trust first. Everything else follows.

The healthcare organizations that will lead the AI era are not those with the most sophisticated models. They are those with the most trusted data. Semantic trust is not a pre-condition to be managed and moved past. It is the architectural property that determines whether every downstream AI investment performs or disappoints—whether predictions are explainable or opaque, whether agents act with confidence or cause harm, whether the flywheel compounds or stalls.

The semantic trust flywheel provides a clear sequence: capture every signal, establish meaning at ingestion, build a continuously validated gold layer, deploy AI on that foundation, activate agentic workflows, and let the outputs feed the next cycle. The technology exists. The standards exist. The regulatory direction is clear. What remains is the organizational will to sequence the investment correctly.

The flywheel starts with trust. Trust starts with semantics. And semantics is an architectural commitment—the most consequential one in healthcare AI today.

Board-level actions to operationalize semantic trust

- **Evaluate your data foundation:** Assess whether your current interoperability infrastructure provides true semantic consistency or just basic API connectivity.
- **Invest in a semantic trust gateway:** Prioritize the architectural layer that normalizes, reconciles, and tracks the provenance of every clinical signal before it reaches your AI models.
- **Scope AI initiatives strategically:** Deploy AI only on clinical domains where semantic trust is established, mitigating operational risk while driving targeted growth.
- **Build a regulatory asset:** Implement provenance tracking now to ensure all AI predictions are fully explainable, safeguarding against future compliance and regulatory scrutiny.

Four things about interoperability and semantic infrastructure worth saying plainly

Interoperability and AI are one program, not two.

The tendency to separate interoperability infrastructure from AI and analytics costs organizations on both fronts. Investment in the Semantic Trust Gateway directly enhances AI capabilities, while AI-generated insights refine signal capture priorities and data quality requirements. Separate governance creates two teams working against each other's assumptions. A unified architectural vision compounds the investment's value.

Compounding returns favor early movers.

Every improvement to semantic infrastructure—refined terminology mapping, accurate patient matching, or new provenance annotation—enhances the quality of all downstream AI models. For example, a single improvement to SNOMED CT mapping for cardiovascular conditions benefits every predictive model, care gap analysis, and agentic workflow touching that concept. This is compounding, not linear, value. Early investors in semantic infrastructure pull further ahead, while late adopters fall progressively behind.

Large language models raise the stakes.

A common argument suggests generative AI reduces the need for reliably coded data, as LLMs can structure unstructured text directly. While partially true about capability, this misses the implication. LLMs can generate structured clinical output, but without ground-truth validation, it lacks provenance: it cannot be audited, its confidence measured, or its errors systematically corrected. Semantic trust makes LLM-generated structure usable in clinical contexts. It's not an alternative to semantic trust—it's its most demanding application.

This is about sequencing, not technology.

The components of a Semantic Trust Gateway—FHIR, SNOMED CT, MPI, provenance tracking, knowledge graph—are established practices. This framework's contribution lies in the correct sequence for assembling and governing them, emphasizing that the sequence is non-negotiable. Deploying AI before establishing the semantic layer leads to heavy spending on model infrastructure while the underlying data silently degrades. The right choice is deliberate architectural sequencing, made as a board-level strategic commitment.

About Brillio

Brillio is a digital technology services company that drives AI-first engineering and design-led experiences for global enterprises. Born digital in 2014, its consulting-led services span Customer Experience, Data & AI, Product Engineering, and Digital Infrastructure. With an industry-leading NPS of 71, Brillio accelerates time to market through its proprietary BrillioOne.ai platform, powered by AI-ready talent with deep domain expertise.

Brillio is the official Digital Transformation Partner and the official Data and AI Services Provider of Atlassian Williams Racing. Brillio partners with leading technology providers including Microsoft, AWS, Google Cloud, Salesforce, Adobe, Databricks, and Snowflake and operates with 6,000+ “Brillians” across 15 global delivery centers. Consistently recognized as a Great Place to Work® since 2021, Brillio blends innovation, talent, and purpose to deliver measurable outcomes for clients and fulfilling careers for employees.



<https://www.brillio.com/>

Contact Us: info@brillio.com

brillio
●●