

Beyond the Algorithm

How memory knowledge graphs can preserve and scale human judgment in non-deterministic enterprise decisions.

Central proposition. Many important enterprise decisions do not have a single objectively correct answer. They depend on incomplete evidence, organizational preferences, precedent, experience and professional judgment. The next generation of enterprise systems must therefore preserve not only facts and outcomes, but also the context and reasoning through which decisions were made.

Author Nish Parikh, Founder and Chief Executive Officer, AgentforgeOS, Inc.

Published October 2026 · version 5

Subject Organizational memory, human judgment and the architecture required to retain both

Abstract

Artificial intelligence has improved the speed with which organizations analyze data, generate content and automate routine work. Yet many of the decisions that create the most value remain resistant to full automation. Forecasting demand for a new apparel product, structuring a client proposal, drafting a legal brief or defining a Statement of Work often involves ambiguous objectives, weak signals, incomplete information and several defensible solutions. In these settings, expertise consists less of applying a fixed rule than of recognizing situations, weighing evidence, understanding preferences and recalling what happened in comparable cases. This paper examines why conventional data systems fail to preserve that form of knowledge, why human judgment remains indispensable, and why memory knowledge graphs provide a promising architecture for capturing organizational context, reasoning and experience without pretending to eliminate uncertainty.

I The most consequential decisions are often not deterministic

For decades, organizations have invested in becoming data-driven. Data warehouses, business-intelligence platforms, forecasting models and optimization engines have transformed operations when the problem is well defined, the objective can be measured and the relevant variables are reasonably stable. These tools are indispensable, but their assumptions do not describe every important decision.

Consider forecasting demand for a new apparel product. Many products may share nearly identical structured attributes: price point, category, color, channel, geography and target segment. Their outcomes can still differ sharply because demand is shaped by interactions among fashion, merchandising, competitor availability, local events, weather, social influence and timing. Each variable may have weak explanatory power in isolation, and some of the most relevant evidence may never enter an enterprise system.

Even the objective can be uncertain. The cost of a stockout is not simply the retail value of the unavailable unit. A customer may wait, substitute, move to another channel, switch brands or abandon the purchase. Inventory carrying cost similarly varies with financing conditions, obsolescence risk, space constraints and the strategic value of availability. The difficulty is therefore deeper than poor data quality: the relationship between action and outcome is partially observable, path-dependent and unstable.

Herbert Simon's theory of bounded rationality challenged the assumption that decision-makers can evaluate every alternative and optimize against complete information. In practice, people operate under limits of time, information and cognitive capacity, and often seek a satisfactory course of action rather than a theoretically perfect one. **[1]**

This is equally true in professional services. A proposal, legal brief or Statement of Work usually has several viable forms. An expert does not discover the one mathematically correct answer. The expert selects and shapes an approach using precedent, client preferences, risk tolerance, organizational capabilities and an understanding of which trade-offs matter in the present situation.

2 Expertise is contextual pattern recognition

Experts are distinguished not merely by access to more information, but by how they organize and interpret it. Naturalistic decision-making research shows that experienced professionals frequently recognize a situation as belonging to a familiar pattern, identify a plausible response and mentally simulate whether it will work. They do not necessarily generate and score an exhaustive list of alternatives.

A senior planner may remember that a customer routinely inflates forecasts before promotions. A proposal leader may know that a particular client responds to phased delivery and distrusts aggressive transformation claims. A lawyer may recall that an apparently favorable precedent failed in practice because of a narrow procedural distinction. Such knowledge is conditional: this action worked for this type of situation, under these constraints, for these reasons.

That conditional structure is difficult to preserve in conventional repositories. A forecast table retains the final number but not the planner's reservations. A proposal retains the selected storyline but not the competing narratives that were rejected. A contract retains the final scope but not the negotiations and preferences that shaped it.

The organization stores artifacts while losing the reasoning that gave those artifacts meaning.

This intuition is not new. Case-based reasoning research established, decades before graph databases existed, that expertise could be represented as a structured library of precedent cases, each retrieved and adapted according to which features made it similar to the situation at hand. **[2]** What has changed is not the underlying logic but the ability to apply

it automatically, continuously and at enterprise scale, across every document an organization produces.

3 Organizational memory is distributed, and fragile

No single person possesses the knowledge required for a complex enterprise decision. Relevant experience is distributed across functions, geographies and levels of seniority. It also resides in email, meetings, presentations, transaction systems, informal conversations and work practices. Organizational-memory research has long treated memory as something retained not only by individuals, but also through structures, processes, culture and external records. [3]

The problem is not that enterprises lack information. They produce more of it than at any earlier point. The problem is fragmentation. Evidence is separated from the decision it informed. Assumptions are not linked to later outcomes. Dissent disappears after consensus is reached. Expertise leaves when experienced employees retire or change roles.

This fragmentation is not only a conceptual loss; it is a measured one. [4]

WHAT THE SURVEY EVIDENCE SHOWS

42%

of the knowledge needed to do a role is held only by the person in it

\$47M

lost per year by the average large US business to inefficient knowledge sharing

5.3 hrs

lost weekly per knowledge worker waiting for or rebuilding what already exists

On these numbers. From an online survey of 1,001 US adults at organizations of 200 or more employees, each with at least five years of experience. The study was commissioned by a vendor, and the \$47M is a model combining expected losses from inefficient information exchange with expected losses during onboarding, not a direct measurement. They indicate a magnitude. They are not precise.

What those figures are not ambiguous about is where the loss concentrates: not in facts that can be looked up, but in reasoning, dissent and context that were never captured anywhere a colleague or successor could find them.

An intern cannot be expected to forecast with the judgment of a professional who has spent twenty years observing products, customers and exceptions. But a memory system can expose the intern to the expert's way of thinking. It can surface comparable cases, explain

which signals senior planners trusted, show which assumptions failed and reveal how the organization balanced risk in similar circumstances.

The objective is not to give every employee the same experience. It is to make the organization's accumulated experience available at the moment of decision.

4 Human heuristics are both capability and risk

Human beings use heuristics because deliberate analysis of every possibility is impossible. These shortcuts reduce cognitive load and permit rapid action. In fast-moving environments, experienced people can sometimes recognize a developing situation and act before an automated system has fully classified it. Pilots, clinicians, emergency responders and operators regularly combine formal procedures with adaptive judgment.

The same heuristics also create predictable errors. Anchoring can make an initial estimate disproportionately influential. Confirmation bias can narrow the evidence considered. Availability and recency can cause a memorable case to outweigh a more representative history. Overconfidence can suppress dissent. The value of an enterprise memory system is therefore not to remove human discretion, but to challenge it constructively.

Recent human-AI teaming research emphasizes that effective performance depends on the design of the combined team. AI can support situation awareness, integrate heterogeneous information and expose the basis for a recommendation, while humans remain responsible for interpreting goals, novelty and context. The National Academies also notes that human biases can both reduce cognitive load and produce error, making bias mitigation a design problem for the human-AI system rather than a reason to exclude human judgment. [5]

A useful system might therefore warn that a recommendation is overly anchored to the previous quarter, that only confirming examples have been retrieved, or that a preferred proposal structure reflects the delivery team's habits rather than the client's demonstrated preferences. Such prompts do not dictate the answer. They broaden the evidence considered and make the reasoning more deliberate.

5 Decision quality cannot be reduced to outcome quality

Many enterprise decisions produce delayed and noisy feedback. A planner may increase inventory before sales rise, but the outcome may also reflect a competitor stockout,

favorable weather or an unexpected campaign. A proposal may be lost for reasons unrelated to its quality. A sound legal strategy may fail because of a novel interpretation. Conversely, a weak decision may benefit from luck.

A decision should be evaluated against what was knowable at the time: the evidence available, the assumptions made, the alternatives considered and the risks accepted.

Capturing only the final result teaches the wrong lesson when causality is uncertain. This distinction also changes how enterprise AI should learn. A binary label, successful or unsuccessful, is often an inadequate representation of judgment. A richer memory preserves the decision process and allows future users to compare both reasoning and outcomes across related situations.

6 The requirements of an organizational memory system

A system designed for judgment-intensive work must preserve more than documents and extracted facts. It must represent the relationships that make facts meaningful and the history through which interpretations change. At minimum, it should connect:

Memory element	What must be preserved
Expectation	What the organization believed should happen, per unit and per period, with its source, version and effective date
Situation	The entities, events, constraints and temporal conditions surrounding a decision
Evidence	Supporting and contradictory sources, with provenance, freshness and confidence
Judgment	Assumptions, preferences, trade-offs, dissent and alternatives considered
Decision	What was chosen, by whom, when and under which authority
Outcome	What subsequently occurred, including uncertainty about causation and measurement
Analogy	Prior situations that are similar in relevant ways, and the dimensions on which they differ

6.1 What memory alone does not solve

The first element in that table is the one most often missing, and it is worth separating from the rest. The other six describe a memory of decisions. The first describes a view of what was supposed to happen, which is a different kind of object and serves a different purpose.

A memory of past decisions answers the question *what did we do last time this happened?* It does not answer the question that comes before it: *is anything happening?*

An organization can hold a complete record of every prior decision and still fail to notice that the plan it is operating under stopped being true five months ago. Precedent is retrieved when someone already knows there is a decision to be made. The more common and more expensive failure is that nobody knows yet. No one decided badly; no one decided at all, because nothing in the system held a view of what should have been occurring and compared it against what was.

This requires a structure alongside memory: a maintained expectation of what should be happening, per product, per account, per period, carrying its own source, version and effective date. Actuals are compared against it continuously, and a gap beyond a declared threshold is what raises a situation in the first place.

Three properties distinguish an expectation from a forecast, and all three matter:

- **It is declared, not inferred.** Someone states what the expectation is and takes responsibility for it. A forecast is a model's output; an expectation is a commitment the organization has made, which is why a gap against it is actionable and a gap against a forecast is often just model error.
- **It is versioned and dated.** When an expectation changes, the prior version and the reason for the change are retained. Without this, a plan that was quietly revised looks like a plan that was always met.
- **It is inspectable.** A person can ask what the expectation says, where it came from and when it last changed, and receive an answer in the same terms they would use themselves. An expectation that can only be examined by reading model internals cannot be argued with, and an expectation nobody can argue with is not one the organization has really adopted.

Memory without expectation is an archive: complete, searchable and silent until questioned. Expectation without memory is an alert: timely and shallow, with nothing to say about what to do next.

Judgment-intensive work needs both, and the link between them is what makes a retrieved precedent relevant rather than merely similar, because the precedent worth surfacing is not the one with matching attributes. It is the one where the gap between expectation and reality had the same shape.

7 Why memory knowledge graphs are a natural fit

A memory knowledge graph is not merely a graph database containing enterprise facts. It is an architecture for representing organizational experience as connected, evolving and attributable knowledge. It can link people, documents, events, decisions, assumptions, preferences and outcomes while retaining the provenance and time associated with each claim, the same requirement that underlies established data-provenance standards such as the W₃C data model for representing who or what influenced a given piece of information.

[6]

This matters because judgment is relational. A forecast adjustment is meaningful because it concerns a particular product, customer, promotion and time period; because it relied on evidence from specific sources; because an expert accepted one risk and rejected another;

and because later events either reinforced or weakened the original reasoning. Removing those links reduces the decision to an isolated number.

Recent GraphRAG research demonstrates the growing interest in using graph structure to organize heterogeneous relational information for retrieval and generation. Surveys published in 2024 and 2025 describe graphs as a means of preserving entity relationships, supporting multi-hop retrieval and assembling more context-aware evidence for language models. [7, 8]

Beyond survey-level interest, direct empirical comparisons support the same conclusion. Microsoft Research's GraphRAG system, evaluated on questions that require synthesizing an entire document set rather than a single passage, produced markedly more comprehensive and more diverse answers than a conventional retrieval baseline. [9] Separately, HippoRAG, a retrieval architecture modeled explicitly on the hippocampal indexing mechanisms of human long-term memory, outperformed conventional and prior graph-based retrieval methods on multi-hop question answering by linking information through a persistent, updatable graph index rather than re-deriving connections at query time. [10] A follow-on architecture, HippoRAG 2, subsequently addressed context loss in that entity-centric design, a useful reminder that this is an active, improving area of research rather than a single settled method. [11]

Both results speak directly to organizational memory: they show that graph structure specifically, not simply better retrieval or larger context windows, is what enables synthesis across many linked pieces of evidence, and that a graph index can be extended with new experience over time rather than rebuilt from scratch.

Other recent work applies this same architecture specifically to decision-relevant memory for AI agents. Zep, for example, builds a temporal knowledge graph that anchors facts and events to when they were true, not only to when they were asserted, a distinction equally relevant to enterprise decisions, where knowing when evidence held is often as important as knowing what it said. [12]

Why a graph, and not a well-organized log

A reasonable objection is that a sufficiently well-tagged decision log or repository could achieve much of the same result without the added complexity of a graph. The distinction matters most at the point of retrieval. A log finds records that match a query in isolation; it does not surface the record that is two or three relationships away from the query but decisively relevant, such as the client preference recorded in a call transcript, linked to a proposal section, linked to a commercial outcome. Graph structure makes that kind of

multi-hop, relationship-driven retrieval a native property of the system rather than something a person must reconstruct by searching several repositories and holding the connections in mind. This is precisely the property the empirical comparisons above are measuring, and it is also what allows the system to answer questions no single document was written to answer.

The stronger enterprise opportunity is to extend this approach from connected facts to connected experience. Instead of retrieving documents that happen to mention the same keywords, the system can retrieve prior situations: the relevant actors, constraints, reasoning, disagreements, decisions and consequences. This is closer to how experts recall precedent and apply judgment.

A memory knowledge graph should not be presented as a mechanism for producing certainty where none exists. Its value is more disciplined: preserving what the organization knew, believed and preferred; exposing how those views changed; and allowing humans and AI systems to reason from a richer institutional history.

Conclusion: from information-rich to memory-rich organizations

The next frontier of enterprise intelligence is not simply better prediction or faster content generation. It is the ability to retain and reuse the reasoning that connects information to action.

Organizations currently preserve transactions and final artifacts far more reliably than they preserve context, judgment and experience. They preserve what was expected even less reliably than that, which is why so many failures are discovered rather than detected. This leaves each generation of employees to reconstruct lessons that the enterprise has already paid to learn. It also constrains AI systems to work from fragments: documents without decision history, outcomes without assumptions and facts without relationships.

Memory knowledge graphs offer a path toward a different kind of enterprise system, one that treats human judgment as an asset to be captured, challenged and shared. Such a system does not seek to make experts obsolete. It allows a junior employee to see how experts approached analogous situations, enables an experienced professional to discover evidence outside their own memory, and helps the organization identify when familiar biases are pushing a decision in the wrong direction.

The organizations that succeed will not be those that remove people from every decision. They will be those that combine analytical systems, generative AI and human expertise within a persistent organizational memory, so that experience compounds rather than disappears.

References

1. Simon, H. A. (1978). *Rational decision-making in business organizations*. Nobel Memorial Lecture. See also Simon, H. A. (1947/1997), *Administrative Behavior*.
2. Aamodt, A., & Plaza, E. (1994). Case-Based Reasoning: Foundational Issues, Methodological Variations, and System Approaches. *AI Communications*, 7(1), 39–59.
3. Walsh, J. P., & Ungson, G. R. (1991). Organizational memory. *Academy of Management Review*, 16(1), 57–91.
4. Panopto & YouGov. (2018). *Workplace Knowledge and Productivity Report*. Online survey of 1,001 US adults at organizations of 200+ employees, each with at least five years of work experience. Vendor-commissioned; the \$47M figure is a modeled estimate combining productivity loss from inefficient information exchange with loss during onboarding, not a direct measurement.
5. National Academies of Sciences, Engineering, and Medicine. (2022). *Human-AI Teaming: State-of-the-Art and Research Needs*. Washington, DC: The National Academies Press.
6. Moreau, L., & Missier, P. (Eds.). (2013). *PROV-DM: The PROV Data Model*. W3C Recommendation.
7. Peng, B., Zhu, Y., Liu, Y., et al. (2024). Graph Retrieval-Augmented Generation: A Survey. arXiv:2408.08921.
8. Han, H., Wang, Y., Shomer, H., et al. (2025). Retrieval-Augmented Generation with Graphs (GraphRAG). arXiv:2501.00309.
9. Edge, D., Trinh, H., Cheng, N., et al. (2024). From Local to Global: A Graph RAG Approach to Query-Focused Summarization. Microsoft Research. arXiv:2404.16130.
10. Gutiérrez, B. J., Shu, Y., Gu, Y., Yasunaga, M., & Su, Y. (2024). HippoRAG: Neurobiologically Inspired Long-Term Memory for Large Language Models. *NeurIPS 2024*.
11. Gutiérrez, B. J., Shu, Y., Qi, W., Zhou, S., & Su, Y. (2025). From RAG to Memory: Non-Parametric Continual Learning for Large Language Models. *ICML 2025*.
12. Rasmussen, P., Paliychuk, P., Beauvais, T., Ryan, J., & Chalef, D. (2025). Zep: A Temporal Knowledge Graph Architecture for Agent Memory. arXiv:2501.13956.

Further reading

- Klein, G. (1998). *Sources of Power: How People Make Decisions*. MIT Press.
- Kahneman, D. (2011). *Thinking, Fast and Slow*. Farrar, Straus and Giroux.
- Nonaka, I., & Takeuchi, H. (1995). *The Knowledge-Creating Company*. Oxford University Press.