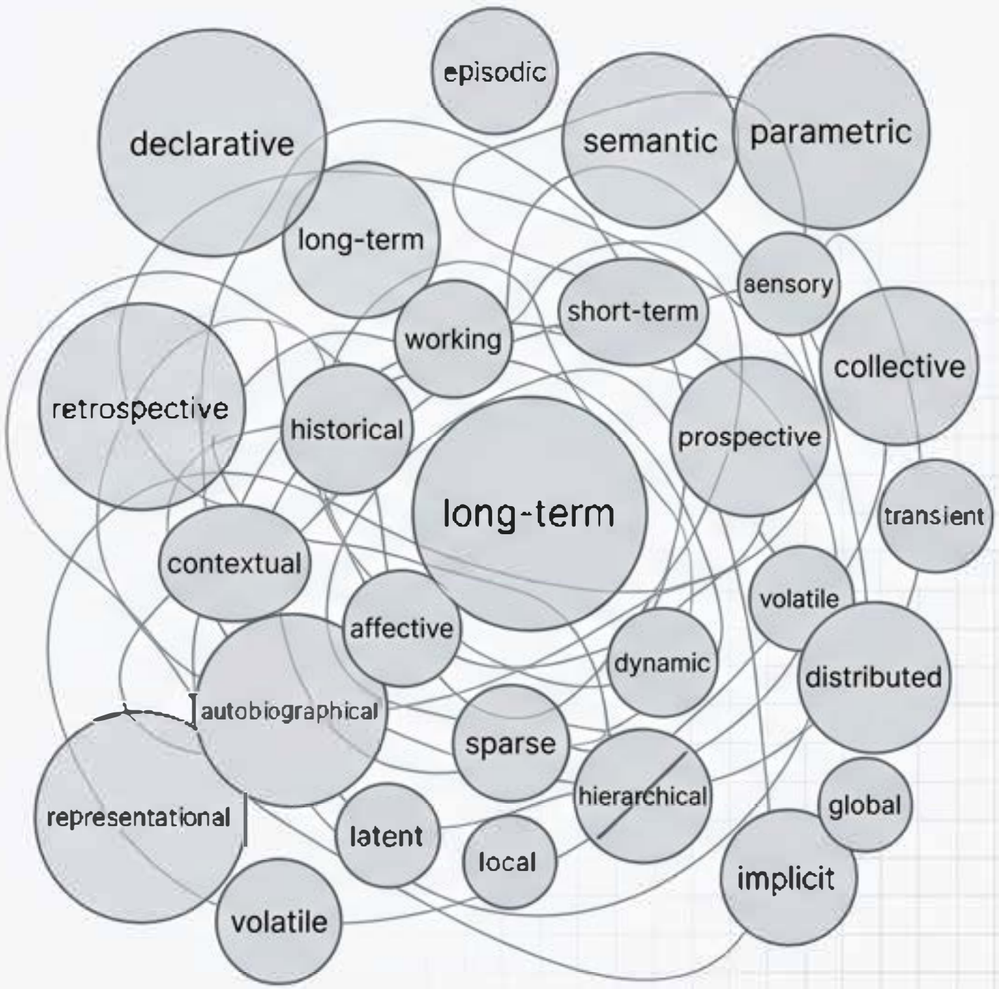
A Blueprint for Memory in Agentic Intelligence

Unifying the Forms, Functions, and Dynamics of AI Agent Memory

The Landscape of Agent Memory is Fragmented

- Research on agent memory is expanding at an unprecedented rate, but the field has become increasingly fragmented.
- Works under the "agent memory" umbrella differ substantially in motivations, implementations, and evaluations.
- A "proliferation of loosely defined memory terminologies" has obscured conceptual clarity.
- Traditional taxonomies like "long/short-term memory" are insufficient to capture the diversity of modern agent memory systems.

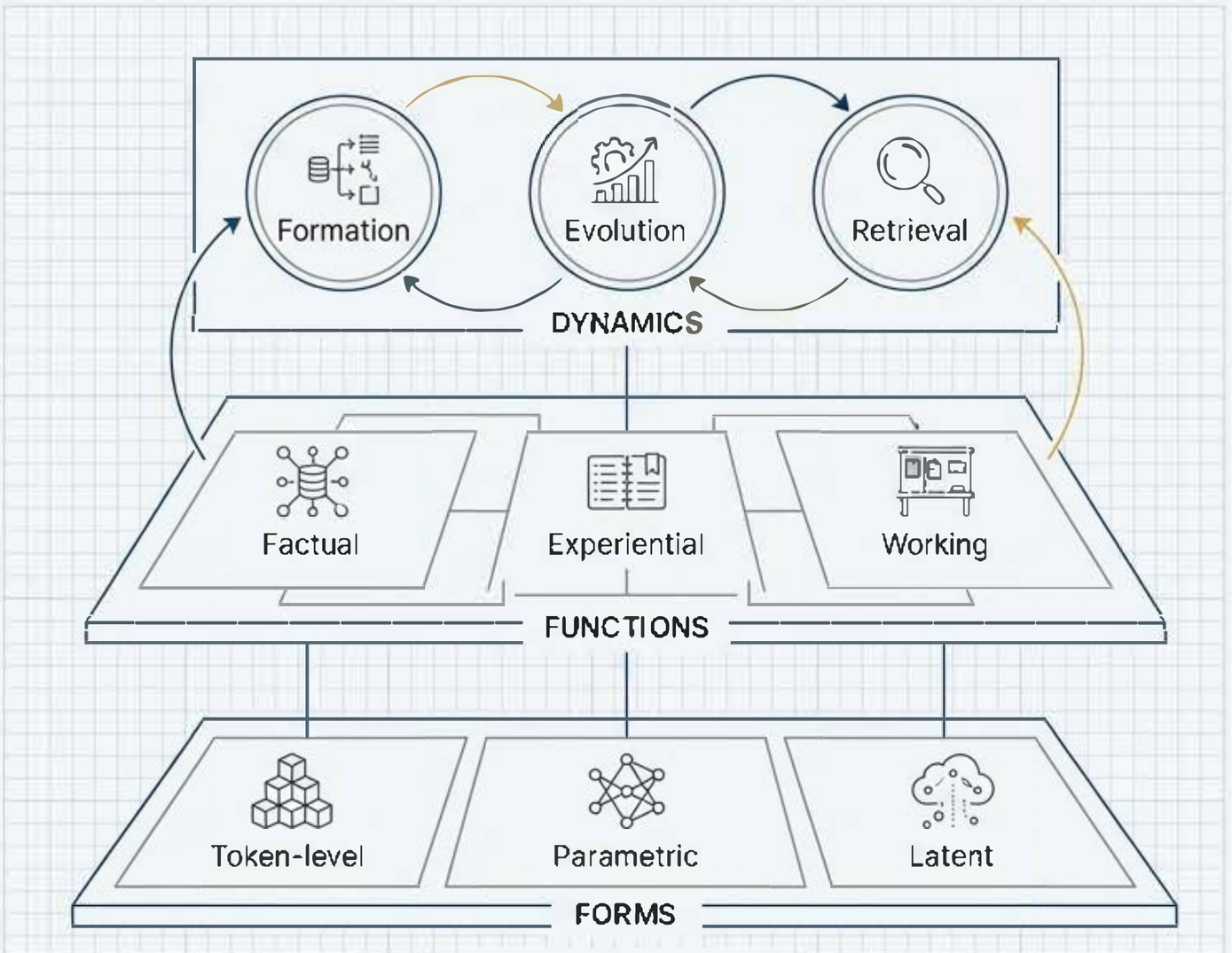
"The explosive growth of memory-related studies... highlights the urgent need for a coherent taxonomy that can unify these emerging concepts."



A Unified Blueprint: Forms, Functions, and Dynamics

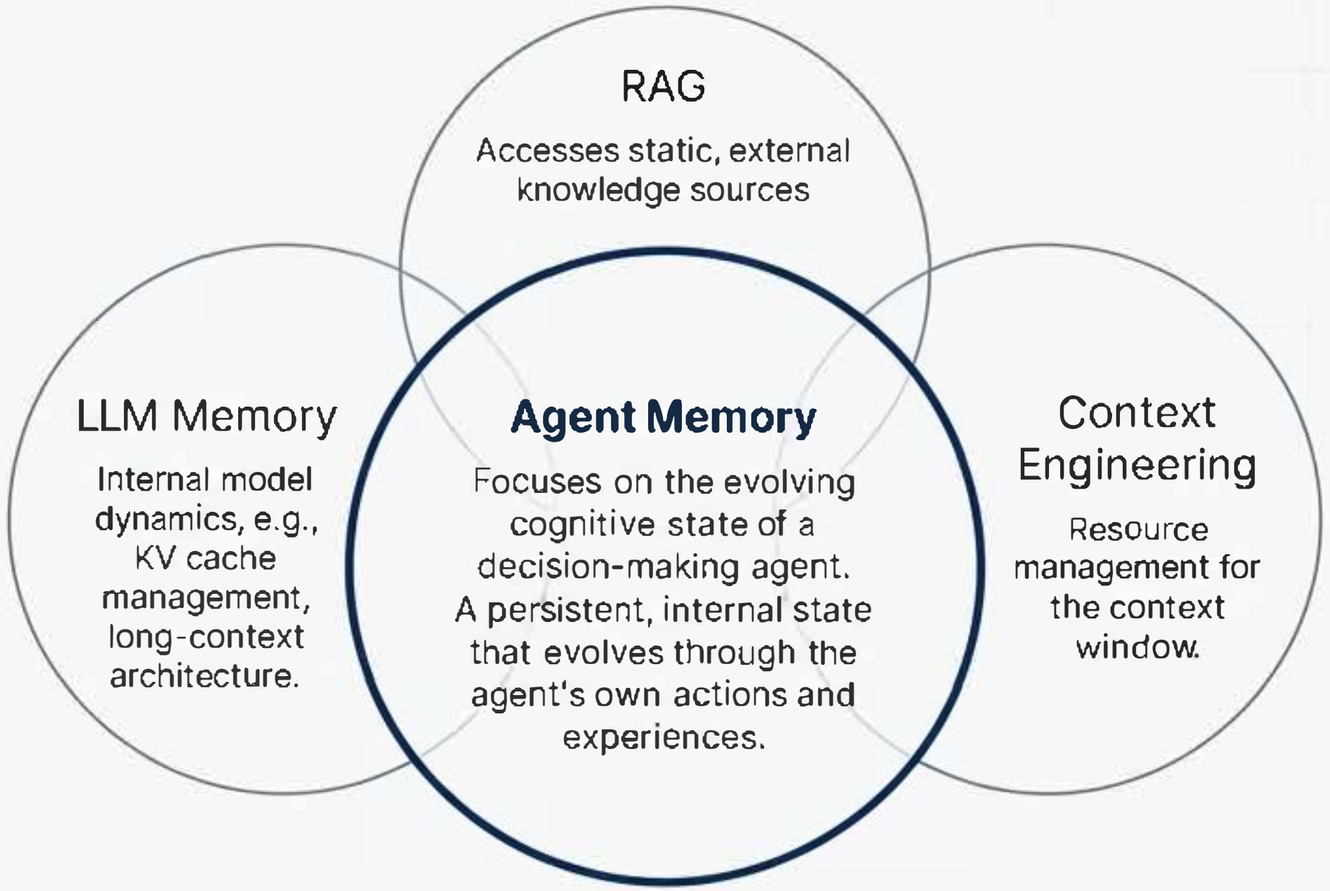
We propose a systematic framework to understand and design agent memory, organized around three core questions:

* **Forms (The WHAT):** What are the architectural and representational forms of memory? What are the building materials?
* **Functions (The WHY):** What roles and purposes does memory serve? Why are we building these structures?
* **Dynamics (The HOW):** How does memory operate, adapt, and evolve over time? How do these systems work?



Master Blueprint

Defining the Scope: What is Agent Memory?



Key Insight: Agent Memory is uniquely characterized by its focus on maintaining a persistent and self-evolving cognitive state that integrates both factual knowledge and experience.

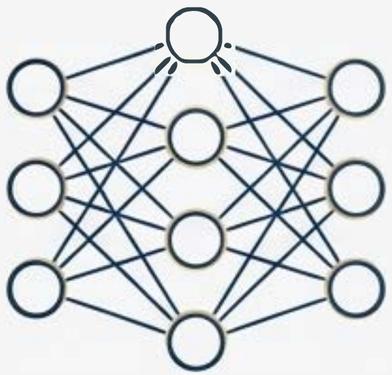
The Materials of Memory (Forms)

Memory is realized in three dominant architectural forms, defined by where and how information is stored.

Token-level Memory

Memory organized as explicit, discrete, and externally inspectable units (e.g., text, visual tokens).

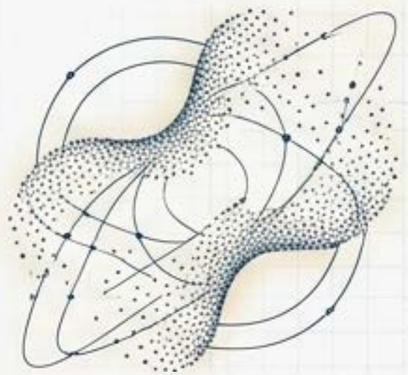
Raw materials like bricks, text files, or data entries.



Parametric Memory

Memory stored within the model's parameters (weights and biases), accessed implicitly during computation.

The ingrained knowledge of a master builder.



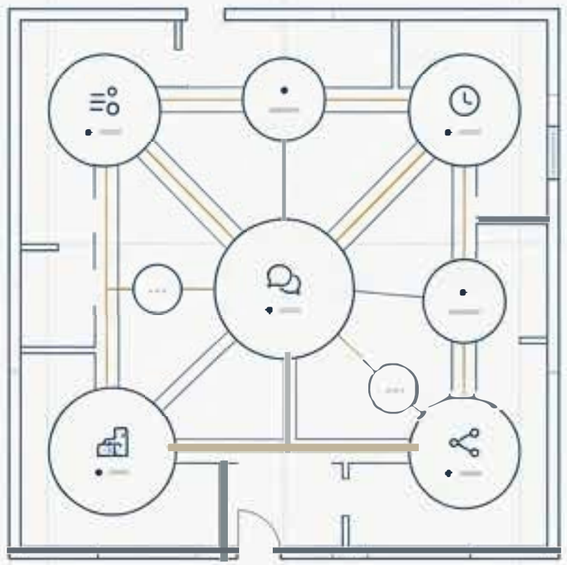
Latent Memory

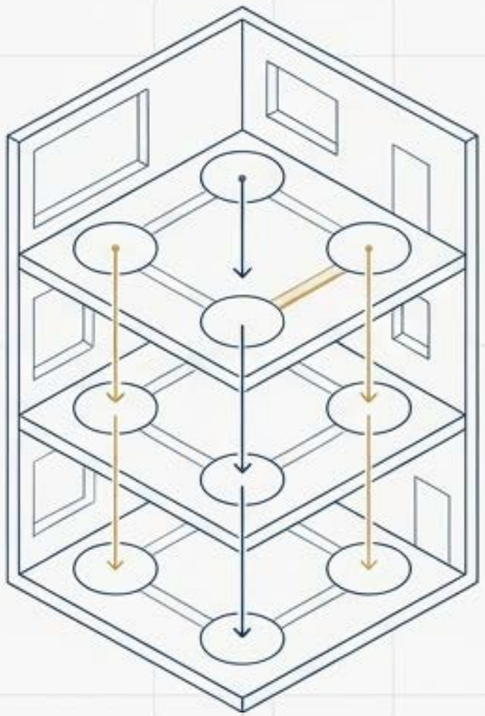
Memory represented in the model's internal hidden states or continuous embeddings.

The working sketch or mental model of the architect.

Structuring the Materials: The Topologies of Token-level Memory

The organization of discrete memory units determines their efficiency and power. We classify them by their structural complexity:





1. Flat Memory (1D)

No explicit inter-unit topology. A sequence or bag of memory units (e.g., dialogue logs, experience pools).

2. Planar Memory (2D)

A single-layer organization, like a graph or tree, connecting memory units.

3. Hierarchical Memory (3D)

Organized across multiple layers with inter-layer links, supporting different levels of abstraction.

The Purpose of Memory (Functions)

Agent memory serves three distinct functional roles, moving beyond simple temporal categories like "long-term" and "short-term."

Factual Memory

The agent's declarative knowledge base. Ensures consistency, coherence, and personalization.

What does the agent know?

Experiential Memory

The agent's procedural and strategic knowledge. Enables continual learning and self-evolution from past successes and failures.

How does the agent improve?

Working Memory

The agent's active, capacity-limited workspace for in-the-moment reasoning and context management.

What is the agent thinking about now?

Deep Dive: How Agents Learn from Experience

Experiential memory is not monolithic. It represents a spectrum of abstraction, transforming raw episodes into reusable capabilities.

Case-based Memory

Stores raw or minimally processed trajectories and solutions. Serves as concrete, verifiable evidence for in-context learning.

Examples:
`Memento`, `Agent KB`, `JARVIS-1`

Strategy-based Memory

Distills transferable reasoning patterns, workflows, and high-level insights from past experiences.

Examples:
`AWM`, `H2R`, `ReasoningBank`

Skill-based Memory

Encapsulates executable procedures, from code snippets to APIs, operationalizing abstract strategies into verifiable actions.

Examples:
`Voyager`, `ToolLLM`, `Alita`

The Operations of Memory (Dynamics)

Memory is a living system, governed by a continuous lifecycle of three core processes.

Memory Evolution

The process of integrating new units into the existing memory base via consolidation, conflict resolution, and forgetting.

How is memory refined?

Memory Formation

The process of transforming raw interaction history and environmental feedback into information-dense, storable memory units.

How is memory extracted?

Memory Base

Memory Retrieval

The process of constructing a context-aware query to access relevant memory content to support a current task.

How is memory used?

Agent Backbone

Selecting the Right Components for the Job

The choice of memory form is not arbitrary; it must align with the agent's task and desired capabilities. Different forms are naturally suited for different applications.

	Memory Form	Key Features	Best Suited Applications
	**Token-level**	Symbolic, addressable, transparent. Swift updates. High interpretability.	**Multi-turn chatbots**, **Personalized agents**, High-stakes domains (legal, medical), **Recommender systems**.
	**Parametric**	Implicit, abstract, generalizable. Slower updates but better performance gain. Prone to catastrophic forgetting.	**Role-playing**, **Reasoning-intensive tasks**, Tasks requiring new fundamental capabilities or stylistic consistency.
	**Latent**	Implicit, human-unreadable (privacy-preserving). Efficient and machine-native. Excellent for modality fusion.	**Multimodal memory**, **On-device or edge deployment**, Low-resource settings.

The Frontiers: Where Do We Build Next?

Our framework reveals several emerging and underdeveloped research frontiers for advancing agentic intelligence:

Automation-Oriented Memory Design: Creating memory systems that can autonomously structure, maintain, and optimize themselves.

Deep Integration with Reinforcement Learning: Using RL to learn optimal policies for memory formation, evolution, and retrieval.

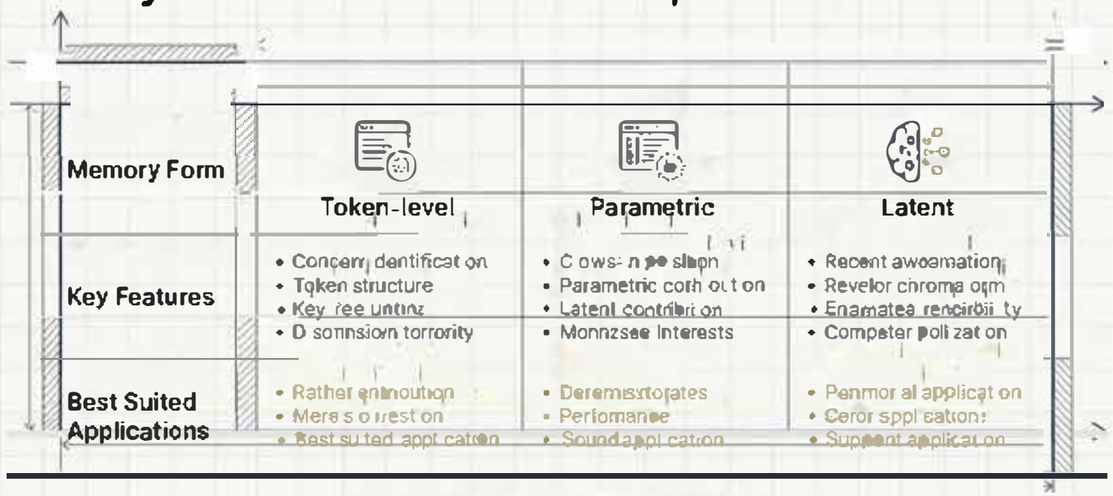
Multimodal Memory: Developing unified memory forms that can seamlessly store and relate information from text, images, audio, and other modalities.

Shared Memory for Multi-Agent Systems: Designing protocols and structures for effective knowledge sharing and collaboration between multiple agents.

Trustworthiness and Memory: Addressing critical issues like privacy, security, and the verifiability of an agent's stored memories and experiences.

Memory as a First-Class Primitive

The future of AGI requires us to move beyond treating memory as an afterthought. It must be a **first-class primitive** in the design of agentic intelligence. The Forms-Functions-Dynamics framework provides the blueprint to build it.



Master Blueprint: Slide 3