Scaling Out Superintelligence

Building an Internet of Cognition for distributed artificial superintelligence.

Vijoy Pandey, Ph.D

V.2 – January 2026

Contents

Executive summary

This paper examines a fundamental shift in AI progress: from scaling the intelligence of individual agentic systems to scaling distributed superintelligence. It argues that today's AI agents, despite their growing capabilities, remain constrained because of semantic isolation. Beyond vertically scaling agents, the next frontier requires new foundational infrastructure for horizontal scaling of intelligence. Such infrastructure would enable shared intent and shared context, accelerate collective innovation, and unlock genuinely emergent capabilities in multi-agent human–AI systems. Together, these systems can address a far broader class of intelligence problems with greater accuracy and robustness.

The paper outlines an Internet of Cognition, a new architectural approach to achieve this transformation and usher in an era of distributed artificial superintelligence. It introduces three necessary components of the architecture: a set of cognition state protocols at latent, compressed, and semantic state granularities (for aligning and coordinating intent), a trusted policy-governed fabric (for sharing knowledge and creating institution-wide context graphs[1]), and cognition engines (for accelerating collective innovation). Together these elements demonstrate how open, interoperable standards are essential for creating multi-agent-human societies that can solve for novel complex problems and discoveries.

1. An agentic life: Full of connection, but without meaning

The advent of powerful foundation models (FMs) has marked a new era in artificial intelligence (AI)—the move from deterministic to non-deterministic, or probabilistic, computing.

In the last few years, it has become evident that FM-based, AI-native applications—especially in the form of agents—are going to disrupt all human work; starting with technology work, to knowledge work (scientific and math), services work, physical work (embodied agents), and even social interactions (agentic avatars). This marks a fundamental platform shift in computing, as we move from cloud native computing to agent native computing.

To solve for this opportunity, frontier model companies have been building more comprehensive and capable agents, driving both solution-space breadth and accuracy by scaling data, compute, and parameters (model size). In other words, up until now, we have only been scaling agents vertically (creating individual geniuses).

1 AI's trillion-dollar opportunity: Context graphs, Jaya Gupta, Ashu Garg, Foundation Capital

Connecting individual agents

For enterprise-grade, or more complex consumer-focused use cases, it became quite evident through the work of Erik Brynjolfsson and team[2] that these individual agents had to be discovered and connected into teams to solve for broader outcomes.

The Internet of Agents[3] vision introduced an open, interoperable way to connect heterogenous agents into multi-agent systems (MAS) for these enterprise-grade use cases (teams of geniuses). Multiple open source projects and standards formed to address this goal in an open manner, including the Model Context Protocol (MCP), the Agent 2 Agent (A2A) protocol, and the AGNTCY project. For more details on this current state of the art, refer to the appendix.

Collective intelligence

Like any other large-scale system, intelligence can be scaled across two vectors including individual (vertical) and collective (horizontal). Current AI agents only utilize one vector, individual scaling of intelligence, which makes the road to superintelligence (or artificial superintelligence, ASI) a long one.[4]

To tackle the full breadth of superintelligence problems with accuracy, in parallel to unlocking the next frontier of model capability through better algorithmic research, we need to leverage horizontal scaling design patterns and unleash distributed artificial superintelligence.

As an example, no singular agent or human has, or can have, the complete institution-wide context graph for an enterprise.

To accelerate our path towards ASI and multi-agent-human societies, we need to create the infrastructure that enables agents, and humans to think together and scale intelligence collectively.

The definition of ASI is often viewed through two lenses, based on what it can do (economic), and how it thinks or works (technical). We lean towards the economic definition, especially because of our belief that to achieve ASI, multi-agent-human societies need to work together in teams to bootstrap the cognitive evolution. These terms and their definitions are explored in the appendix.

Let's dive into why this is critical, what collective scaling means, and what advantages it provides in our journey towards ASI. This paper first looks at the significant shifts in the evolution of human intelligence and draw parallels to where we are today for ASI.

You can also go directly to the Internet of Cognition architecture for distributed superintelligence, or look at some examples and real-world use cases of why this is needed.

2 Future of Work with AI Agents, Social and Language Technologies Lab, Stanford

3 The original vision paper, along with several enterprise-grade multi-agent use cases.

4 Recent predictions for achieving ASI range anywhere between 5 to 20 years. See section 5

2. Paradigm shifts in scaling human intelligence

The history of human intelligence can be characterized by two distinct, transformative phases of capability. There was a long period of individual intelligence scaling (Shift 1), followed by an explosive cognitive evolution[5] driven by semantic communication, collective knowledge, and shared belief systems (Shift 2).

This exact trajectory is replaying out in silicon.

Shift 1: Scaling individual intelligence

For most of human history, from roughly 300,000 to about 70,000 years ago, intelligence was a solitary and localized scaling phenomenon.

While individuals became smarter, developing and using better tools, communicating symbolically through simple patterns and colors, and executing simple plans like creating basic clothing and shelter[6], they lacked the mechanisms for cumulative and collective evolution, often referred to as the **ratchet effect**[7]. Innovation occurred frequently but disappeared with the innovator or remained contained within small, isolated groups due to the inability to communicate complex, abstract concepts at large.

Shift 2: The cognitive evolution

Around 70,000 years ago, a fundamental transformation occurred. Several breakthroughs in semantic communication constructs–sentences, grammar, and recursive language, allowed humans to achieve three critical things they couldn't do earlier:

1. **Shared intent[8]:** The ability to exchange goals and intentions, coordinate actions and rationale, and to be able to view others as cooperative individuals.
2. **Shared context:** Enable the ratchet effect, allowing innovations and technologies to accumulate over time, growing in complexity far beyond what any single individual could invent during a lifetime.
3. **Collective innovation:** Collectively reason and invent tools, concepts, structures, and guardrails that do not yet exist on this shared context base.

5 Sapiens: A Brief History of Humankind, Yuval Noah Harari, Harper.

6 The Evolution of Human Intelligence, on In Our Time, with Melvyn Bragg, BBC

7 Cultural Transmission, Michael Tomasello, Max Planck Institute for Evolutionary Anthropology. More colloquially referred to as standing on the shoulder of giants,

8 Shared Intentionality, M. Tomasello, Open Encyclopedia of Cognitive Science

3. Paradigm shifts in scaling artificial superintelligence

The defining limitation of AI agents today is isolation.

An AI agent can autonomously perceive, reason, act, and use tools[9]. They are powerful *individual* entities, capable of remarkable tool use and problem-solving. They work alone or in small, designated teams for specific workflows and improvements take place within those boundaries. They are operating in the equivalence of individual scaling, like those early humans.

We are just beginning to connect multiple agents together in teams of individual geniuses, but agents today cannot "think" together. There is no coordination for shared intent. There is no creation of shared memory or beliefs that improve and inform future interactions, decisions, and innovations—in other words, no ratchet effect.

For example, when a sales agent figures out how to handle a complex pricing negotiation, that insight stays isolated. Other agents—even in the same organization—start from scratch on similar problems. The sales agent's breakthrough doesn't inform the forecasting agent.

In another canonical example, a hotel booking agent can only work on goals, constraints, and context that is relevant for booking rooms, e.g., find me a room near this beach within a certain price point. An airline agent can only work on constraints around airports to fly out from-to, number of hops, and price. A car rental agent only knows constraints about availability, price, and classes of cars. All the agents know the customers historical context and preferences individually, but do not know how to book an end-to-end vacation for the customer.

See further examples in section 4 below.

Enabling Shift 2: The cognitive evolution for agentic computing

For AI agents to move from Shift 1 (individual, vertical) to Shift 2 (collective, horizontal) scale, they require breakthroughs that have parallels to human intelligence:

1. **Shared intent:** Agents need to align on common objectives and goals and coordinate decisions through semantic meaning, not just message passing. Cognition state protocols let agents understand what they're collectively solving for and negotiate trade-offs to get there.
2. **Shared context:** Emergent multi-agent-human collective context graphs, memories, ontologies, knowledge graphs and beliefs where insights compound over time. When one agentic system solves a problem, that knowledge becomes available across the system. This is the ratchet effect —progress that doesn't reset with each interaction. The distributed cognition fabric provides this shared context layer.
3. **Collective innovation:** To reason together and invent solutions that don't exist yet, within safe boundaries. Cognition engines act as either accelerators that speed up collective

9 Developments in AI Agents: Q1 2025 Landscape Analysis, The Science of ML and AI

problem-solving, or guardrails[10] that enforce compliance, safety, and constraints.

Now let's look at a few real-world use cases if the full vision of the Internet of Cognition stack was realized, ushering in the agentic cognitive evolution.

4. Example: Distributed superintelligence in practice

In the simplest canonical example above, of three agents booking an end to end trip for a customer, the rental car company agent, the airline agent, and the hotel chain agent will semantically coordinate on a singular shared intent of booking a trip for the customer to a beach destination, work within the shared context of the customers hotel type, car type, and flight histories, and collectively solve for this trip being the best outcome for a given cost, during the designated vacation period.

For a more infrastructure-related use case, consider deploying a first-of-its-kind low-Earth-orbit satellite network. The team: Prometheus[11] agent from Outshift by Cisco (network configuration expert), Themis[12] agent from Mythos Corp in Japan (security and compliance specifics for their environment), and Mythos Corp's chief architect who can make hard business decisions.

No single agent can solve this alone. No agent working with a human in traditional request-response mode can either. It requires collective intelligence across agents and humans.

Today's agent infrastructure handles discovery, access control, and message passing. That's necessary but highly insufficient.

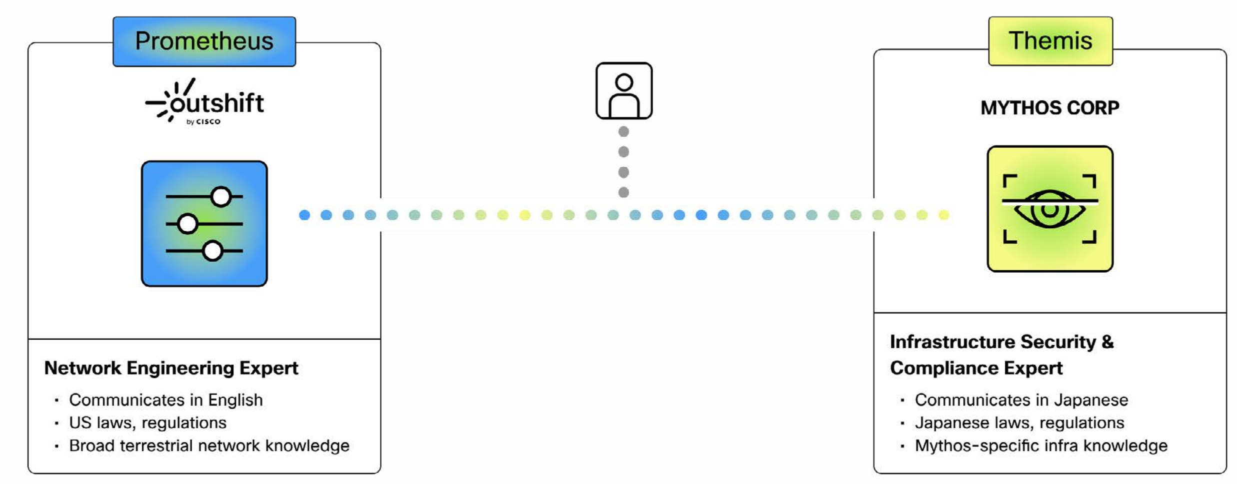
The problem: Prometheus communicates in English, Themis in Japanese, the human operators in both. US and Japanese regulations differ significantly, network design patterns differ significantly, cost structure specifics for the Japanese market are very different from the US market. Outshift's institutional knowledge of terrestrial networks, Mythos Corp's operational specifics, and the chief architect's understanding of business priorities have minimal overlap. This multi-agent-human team can't meaningfully and semantically collaborate on a novel cross-domain problem or even converge on a common objective.

---

10 How Microsoft thinks about AGI, Dwarkesh Patel podcast with Satya Nadella

11 Prometheus the agent, in a real-world use case: Outshift and Swisscom collaborated on agentic network configuration management. The example itself is an illustrative one.

12 Prometheus: creator, inventor. Themis: upholding law and order.



What they need:

1. **Shared intent:** Align on deploying a high-throughput LEO network safely, securely, on time, within budget–with human operators able to inject business decisions throughout
2. **Shared context:** Access to proven network topologies from across Outshift's customer base, Mythos environment specifics, current inventory, and the business constraints that exists with the chief architect
3. **Collective innovation:** Determine which terrestrial network principles apply to low-orbit systems, where to innovate, and when humans need to make judgment calls on trade-offs

The Internet of Cognition makes this work. Cognition state protocols let agents and humans negotiate common goals despite language and regulatory differences. The cognition fabric captures and reconciles institutional knowledge from both organizations plus business knowledge from the chief architect. Cognition engines accelerate joint problem-solving while ensuring compliance with both US and Japanese requirements.

This is distributed artificial superintelligence: specialized agents and humans forming a cohesive intelligent system where collective reasoning emerges from semantic collaboration.

5. An architecture for the agentic cognitive evolution (Shift 2)

The state of the art in agentic collaboration is solving for MAS from heterogeneous vendors, organization and frameworks to be able to simply syntactically talk to each other.

But as we started deploying MAS in the real world[13] [14], we realized that the fundamental challenge is not in the syntactic or framework incompatibilities of multi-agent systems, but in

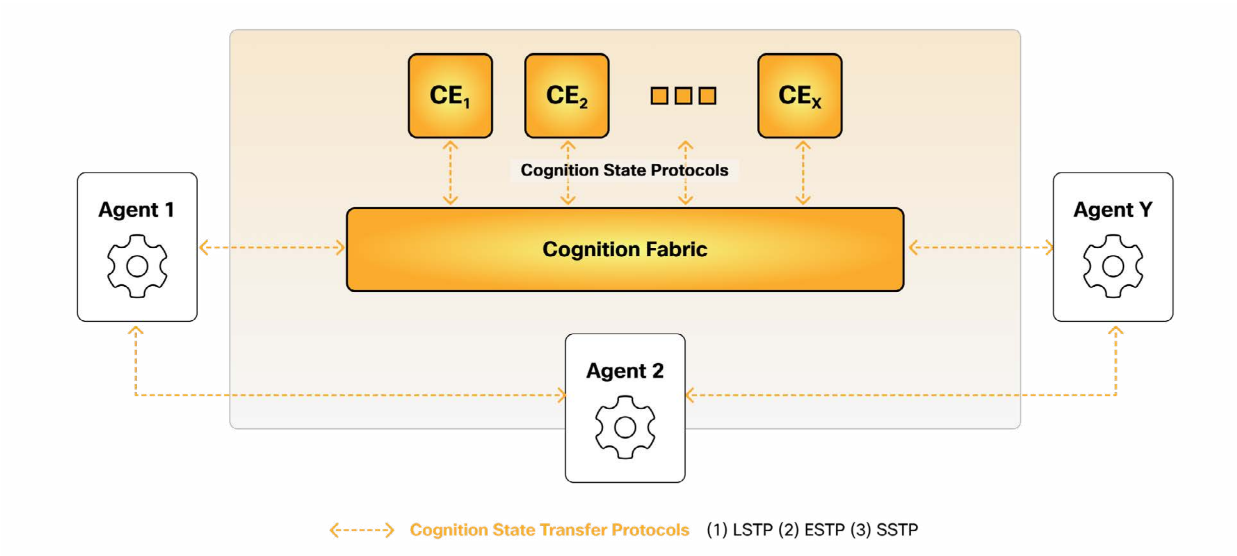
13 Community AI Platform Engineer, https://cnoe-io.github.io/ai-platform-engineering/

14 Autonomous networks, fewer outages: How Swisscom and Outshift are rethinking network change management, through the Agentic Network Validator

the deeper semantic, cognitive, and knowledge layers[15].

It's like allowing a Mandarin-speaking Android user to send text messages to a German-speaking iPhone user on iMessages. The frameworks can connect, minimally, but the users cannot semantically collaborate.

To usher in true collective scaling through an Internet of Cognition, we must build an architecture that supports:



1. **(Shared intent) Cognition State Protocols:** The foundational set of protocols perform five functions: (1) grounding, (2) discovery, (3) resolution, (4) coordination, and (5) negotiation. They define how agents represent and interpret the meaning of exchanged messages; synthesize and manage collective memory, understanding, and knowledge; minimize coherence problems; resolve between local and collective (global) preferences, values, biases and guardrails.

This set of protocols operates at different levels of granularities based on the openness and capabilities of exchange between the agents of a MAS, and the cognition engines at play. We currently identified at least three classes of such protocols:

> **Latent State Transfer Protocol (LSTP):** Class of protocols for inference continuity across models. Allows for high-fidelity transport of raw latent state (activations and internal world-model information), ensuring that the reasoning trajectory is preserved across the endpoints, especially deployed in local clusters with the goal of a unified execution environment.
>
> **Compressed State Transfer Protocol (CSTP):** Class of protocols for invariant alignment across endpoints. Allows for the transfer of compressed and abstracted feature representations, rather than the raw hidden states, allowing agents to achieve consensus without the overhead of the

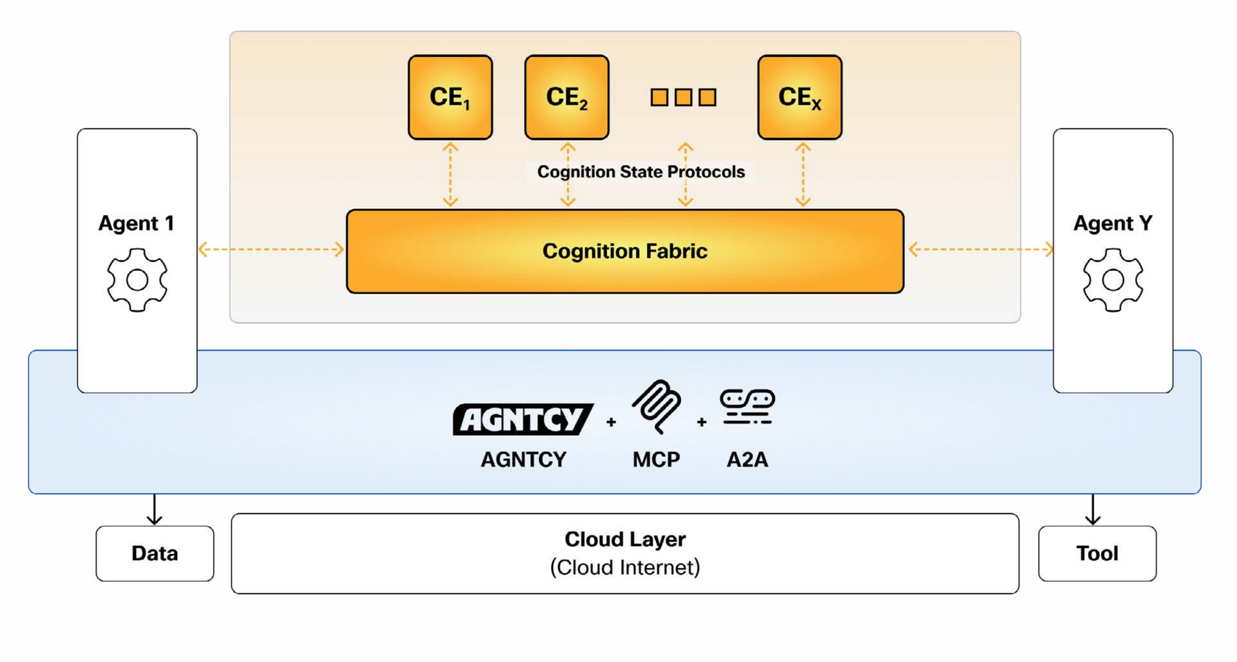
15 A Layered Protocol Architecture for the Internet of Agents, https://arxiv.org/abs/2511.19699

full tensor space. This is particularly useful for low-bandwidth Edge or Wide Area Network (WAN) environments.

Semantic State Transfer Protocol (SSTP): Class of protocols for cross-vendor cross-model cross-agent strategic coordination and human-in-the-loop applications. It flattens the continuous vector spaces of latent and representational models into symbolic or semantic primitives, labels, or formal logic. This provides human-auditable, policy-governed, and model-agnostic handshake for top-level strategic agreements and decisions to move collective innovation forward.

2. **(Shared context) Cognition fabric:** A trusted distributed policy-governed mesh that supports storage, retrieval, modification, i.e., versioned updates for multi-agent-human context graphs, memory and knowledge exchanged via SSTPs, and also the tensor state, latent state, and semantic state as exchanged via LSTP or CSTP. The key is to provide mechanisms to reconcile state information, but also create emergent context information, and build a consistent worldview with granular enterprise-level policies in place.
3. **(Collective innovation) Cognition engines:** A set of cognition engines (CAX in the figure) for cognition functions that are enterprise-grade. There are two types of cognition engines: cognitive amplifiers (COGs) that provide privacy-preserving and secure collective reasoning and exploration assistance; and guardrail technologies (GATs) for security, cost, and compliance[16].

These would all work with the current state of the art infrastructure for cloud computing and agentic computing, extending what's already been deployed rather than replacing it as shown below.

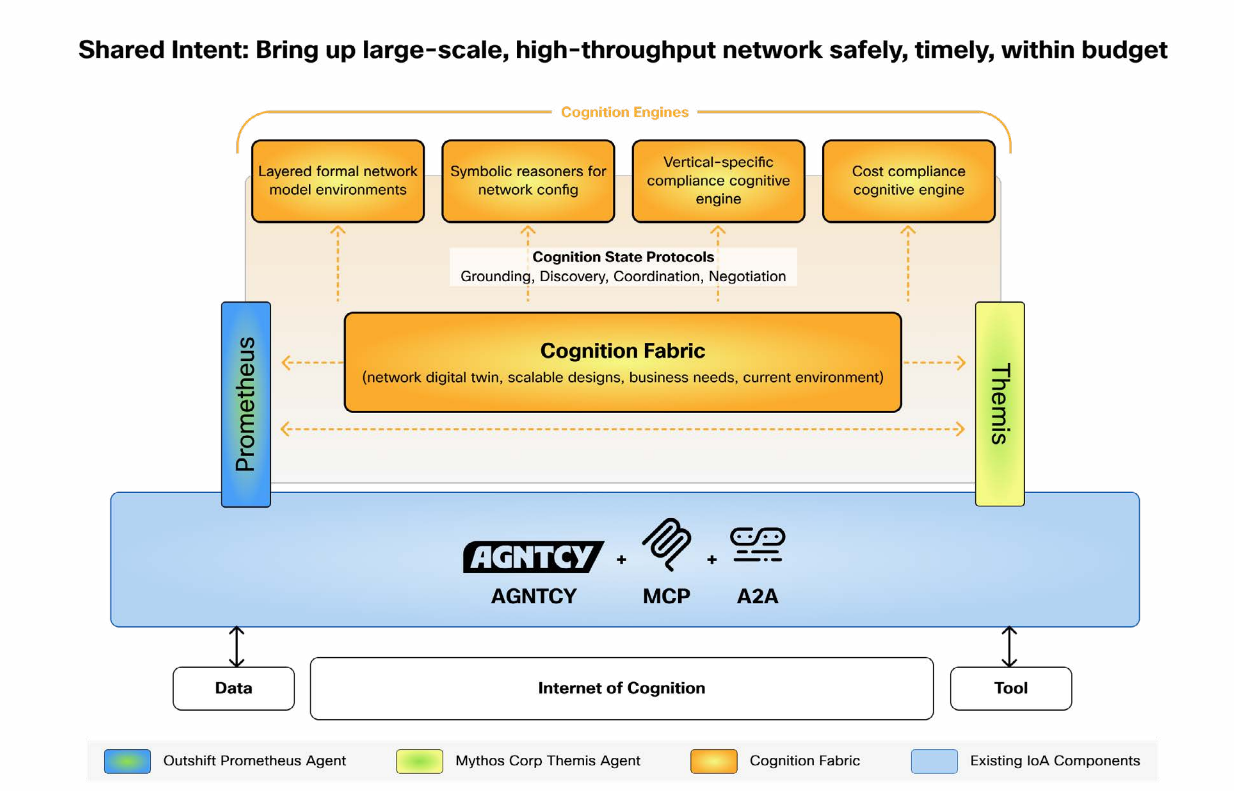


16 "AI should either be guardian angel technologies (GATs) or cognitive amplifiers (COGs)", Raj Reddy, Turing Award Winner

If we architect and build this foundational infrastructure for distributed superintelligence we will create multi-agent-human systems with emergent collective intelligence that usher in the agentic cognitive evolution and accelerate our path towards ASI.

6. Distributed superintelligence apps (multi-agent knowledge workers)

Putting it all together, let's revisit our previous low-earth-orbit network example that the Prometheus and Themis agents are trying to configure for Mythos Corp, and what a possible distributed superintelligence deployment would look like for that use case.



The Outshift Prometheus network engineering expert agent, and the Mythos Corp Themis security and compliance agent agree and coordinate on a shared intent through cognition state protocols. These protocols additionally allow these agents to discover missing information, resolve conflicting information, and, consequently, also update the cognitive memory fabric. Furthermore, these protocols help coordinate and negotiate not just amongst themselves but with the four cognition engines in this deployment.

The cognition fabric provides shared context in the form of working memory, beliefs and design patterns, ontologies like formal network models, and specifics about the Mythos Corp networks, application set, broader infrastructure environment, business goals and service level objectives.

The cognition engines help accelerate collective innovation by providing reasoning and formal modeling capabilities. They also provide guardrails in the form of cost, security, and regulation compliance, where innovation can take place rapidly and safely.

7. Conclusion

Scaling out superintelligence is a call to action. The next era of computation demands a unified commitment to building the infrastructure for distributed superintelligence in an open and interoperable manner.

We urge organizations to:

- Adopt open cognition state protocols for sharing intent and semantic collaboration.
- Explore a distributed cognition fabric for trusted, policy-governed shared memory and knowledge, and to enable net-new emergent behaviors from multi-agent systems.
- Deploy cognition engines for enterprise-grade accelerators and guardrails for collective innovation.

Together, we can create a multi-agent-human future where agents and humans reason, innovate, and solve the world's most complex and novel problems as one.

Appendix

Definitions: Artificial superintelligence (ASI), timelines

The definition of superintelligence or artificial superintelligence (ASI) is often viewed through two fundamentally different lenses, depending on whether you're measuring impact or capability:

1. **Economic:** ASI means that autonomous systems can perform essentially all economically valuable work better than humans, without human oversight. This is how OpenAI's five-level framework[17] and DeepMind[18] position it - it's about what the system can replace or accomplish in economic terms.
2. **Technical:** ASI means that systems can reason towards objectives and generalize far beyond their training data - capable of making genuinely novel scientific discoveries autonomously. This is closer to how researchers like Yann LeCun[19], Ilya Sutskever, and Demis Hassabis think about it - it's about cognitive architecture and capability.

We lean towards the economic definition, especially because of our belief that to achieve ASI, multi-agent-human societies need to work together in teams to bootstrap the cognitive evolution. These definitions are inter-dependent though, since to achieve the economic goal, the technical definition also needs to come true.

The era of vertically scaling existing architectures is hitting an asymptote[20]. With industry pioneers like Ilya Sutskever declaring the age of scaling hitting a plateau, and Sara Hooker[21] exposing the diminishing returns of monolithic (vertical) scale, it is evident that brute-force data, compute, parameter expansion is no longer sufficient.

As a result, the time horizon predictions for ASI have started shifting to the right, with some like Ilya Sutskever stating that we are back into the Age of Research[22], and Yann LeCun and Demis Hassabis claiming that we need a couple of different breakthroughs beyond LLMs to make ASI viable. The current timeline for ASI ranges anywhere between 5 – 20 years[23]. This approach of building towards distributed superintelligence utilizing an Internet of Cognition is the architectural lever to accelerate this timeline.

---

17 OpenAI Scale Ranks Progress Toward 'Human-Level' Problem Solving, Rachel Metz, Bloomberg

18 15 Demis Hassabis: Future of AI, Simulating Reality, Physics and Video Games, Lex Fridman Podcast

19 A Path Towards Autonomous Machine Intelligence, Yann LeCun, Meta FAIR, OpenReview

20 Frontier Language Model Intelligence, Over Time, Artificial Analysis

21 On the slow death of scaling, Sara Hooker, SSRN

22 We're moving from the age of scaling to the age of research, Ilya Sutskever, Dwarkesh Patel podcast.

23 Here are some recent prediction ranges: Ilya Sutskever: 5-20 years, Yann LeCun: 10-15 years, Demis Hassabis: 5-10 years

Definitions: Structure of communication and language

We are inspired by general human language and communication theory[24] to identify three key levels of required agentic communication: syntactic → semantic → cognitive. Additionally, the cognitive layer accumulates and works on data → insights → knowledge and wisdom. This hierarchy can be best described by the following tables.

Structured communication

		In language / humans	In computing / AI
Syntactic	Rules for arranging, representing symbols. Structure of the message	grammar e.g., verb before / after noun	schema and frameworks (e.g., csv)
Semantic	Meaning behind the symbols What messages mean, align interpretability	concepts, relations, intent	meaning representations and models
Cognitive	Interpretation and reasoning, within context Reasoning on the message	decision-making	planning, reasoning, inferencing

Working on

	description	example
Data	raw symbols or measurements	dog, run
Information	data + syntax that organizes it, making it interpretable	noun: dog, verb: run
Knowledge	information + semantics that ties it to real-world meaning or context	Understanding that dog refers to a mammal and run is an action that they perform, and what their relationship is
Wisdom	knowledge + principled ethical judgment in context of consequences	A dog needs to run in open spaces otherwise it suffers

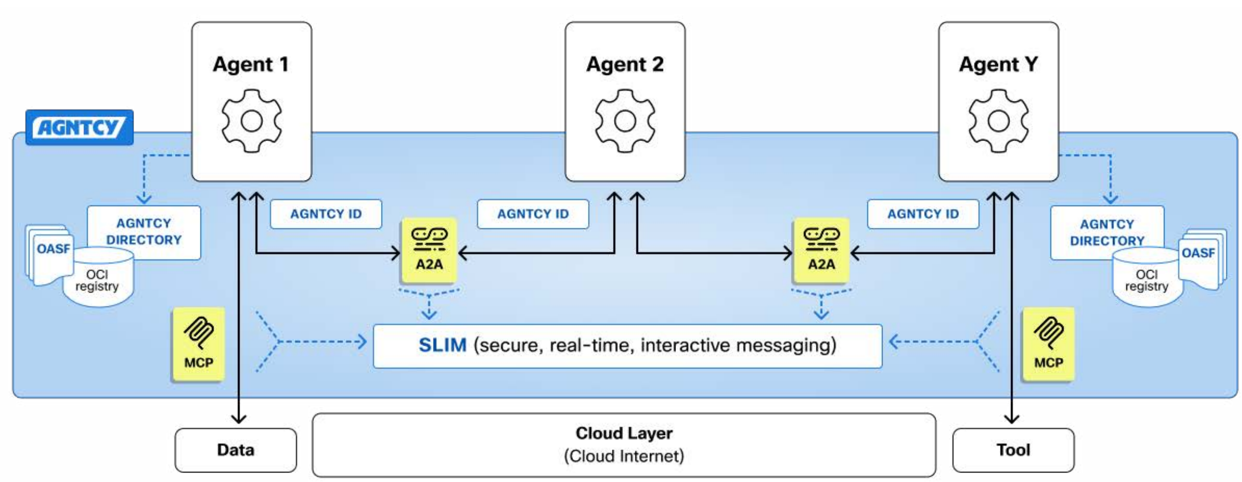
24 Major levels of linguistic structure, Wikimedia.

Landscape: Current state of the art

The Internet of Agents laid out the vision for open, interoperable agent collaboration across vendors, frameworks and organizational boundaries. Multiple projects have formed to address this in an open manner.

> **Model Context Protocol (MCP)** enables communication between an agent and the tools it accesses. Now part of the Linux Foundation-hosted Agent AI Foundation, MCP enriches context for individual agents through standardized tool calling and resource access.
>
> **Agent2Agent (A2A) protocol** enables communication between agents. Donated to the Linux Foundation in June 2025, A2A defines industry-standard patterns for agent-to-agent interaction.
>
> **AGNTCY project** provides discovery, identity, messaging and observability components for the Internet of Agents. It works across protocols including MCP and A2A, offering the Open Agent Schema Framework (OASF) for describing agent capabilities and an Agent Directory for publishing and discovering agents.

Cisco is a launch member of the A2A Project, the AGNTCY Project, and the Agentic AI Foundation (AAIF), all three hosted by the Linux Foundation[25].



Together, these projects solve for the four phases of multi-agent software development and deployment:

1. **Discovery:** A primary obstacle is the absence of a standardized mechanism for developers to locate, verify, and trust the quality and compatibility of available AI agents for a given workflow.

25 Model Context Protocol (MCP), https://modelcontextprotocol.io/; Agent AI Foundation, https://aaif.io/; Agent2Agent Protocol (A2A), https://github.com/a2aproject/A2A; The AGNTCY open source collective, https://agntcy.org

To solve the discovery problem, AGNTCY provides the Open Agent Schema Framework (OASF), a vendor-agnostic standard for describing agent capabilities, and an Agent Directory, which functions as a decentralized registry for publishing and discovering these agents.

2. **Identity:** Provides cryptographically verifiable identity and access control to ensure agents can act securely across organizational boundaries.
3. **Messaging:** The Secure Low-latency Interactive Messaging (SLIM) messaging layer allows for quantum-safe real-time multi-modal communication for running multi-agent systems at scale. The Model Context Protocol (MCP) and the Agent2Agent (A2A) Protocol are both supported for tool and agent communication, respectively.
4. **Observability:** Measuring the effectiveness of multi-agent workflows through extensions to OpenTelemetry and applications for evaluation and experience monitoring.

outshift.com ↗

@ outshiftbycisco