The future of Knowledge is *spatial*

Rebuilding web infrastructure for agentic, non-deterministic intelligence.

Most enterprise AI pilots fail because they discard the relationships inside corporate data. Without this structure, agentic LLMs hallucinate, break compliance rules, and waste budgets

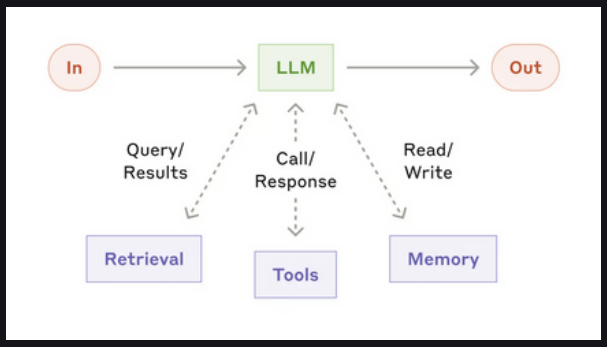
Overview

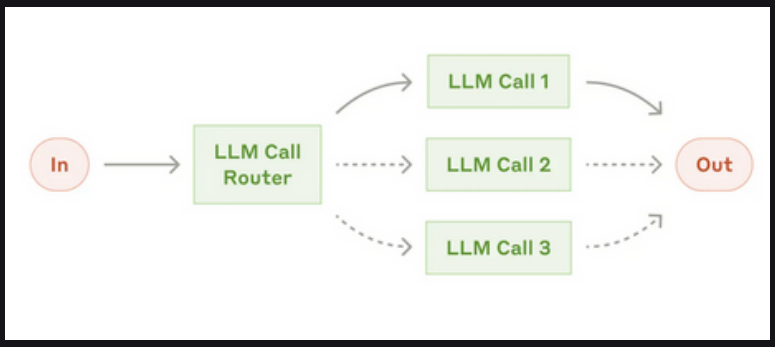
The internet is undergoing a paradigm shift towards a **non-deterministic, agentic** web.

LLMs have an inability to understand and utilise **spatial** and **structural** components of information. The data is often **compressed** into text before passing into the LLMs, losing the key details. This results in suboptimal performance for AI agents and a low return on investment for enterprises wanting to adapt their operations for the agentic age.

Knowledge has components more than just text. Large Language Models (LLMs) do not understand this. They just see autoregressive sequential strings of data. Only the **semantic** understanding is kept, losing both the **implicit** and **explicit spatial representation** found in information.

AI Agents are LLMs with **internal states** (memory / context) and a **set of tools** that it can call (database operations, functions, external API calls) (Google Cloud, 2025). **RAG** (Retrieval Augmented Generation) defines getting **data**, **transforming** it, then **passing into** an LLM for relevant output (Google Cloud, 2024).





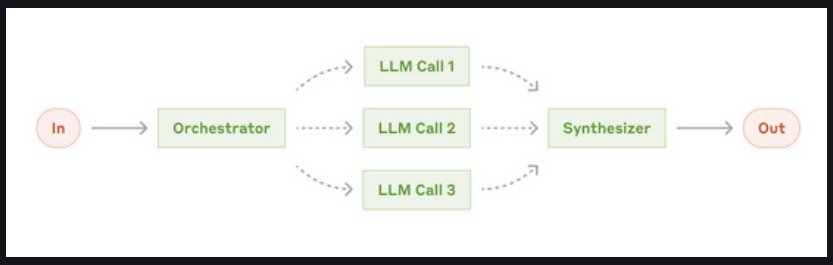


Figure 1: Agentic workflow patterns (Anthropic.com, 2024)

As LLMs become better at reasoning, the input fed into these systems becomes increasingly important — known as the '**context window**'. Research has found that context windows can be **poisoned** by irrelevant data (Anthropic.com, 2025). For traditional RAG applications, the data is flattened into unstructured text strings after retrieval.

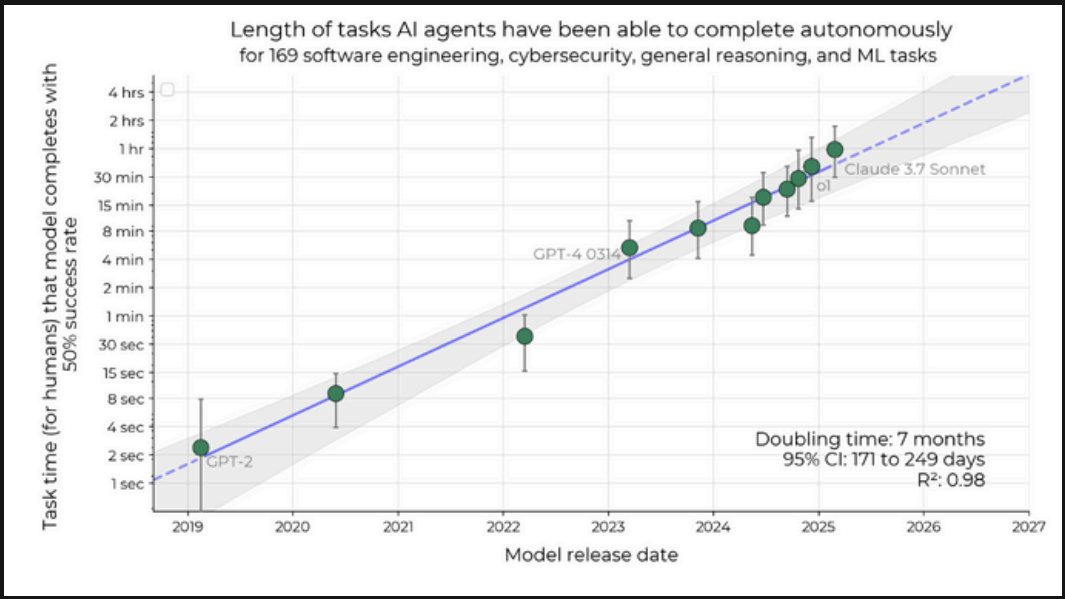


Figure 2: AI time-horizon is extending. (Kwa et al., 2025).

In scope: the context layer for enterprise LLM apps (representation, retrieval, evaluation).
Out of scope: model training, model choice, and end-user UX.

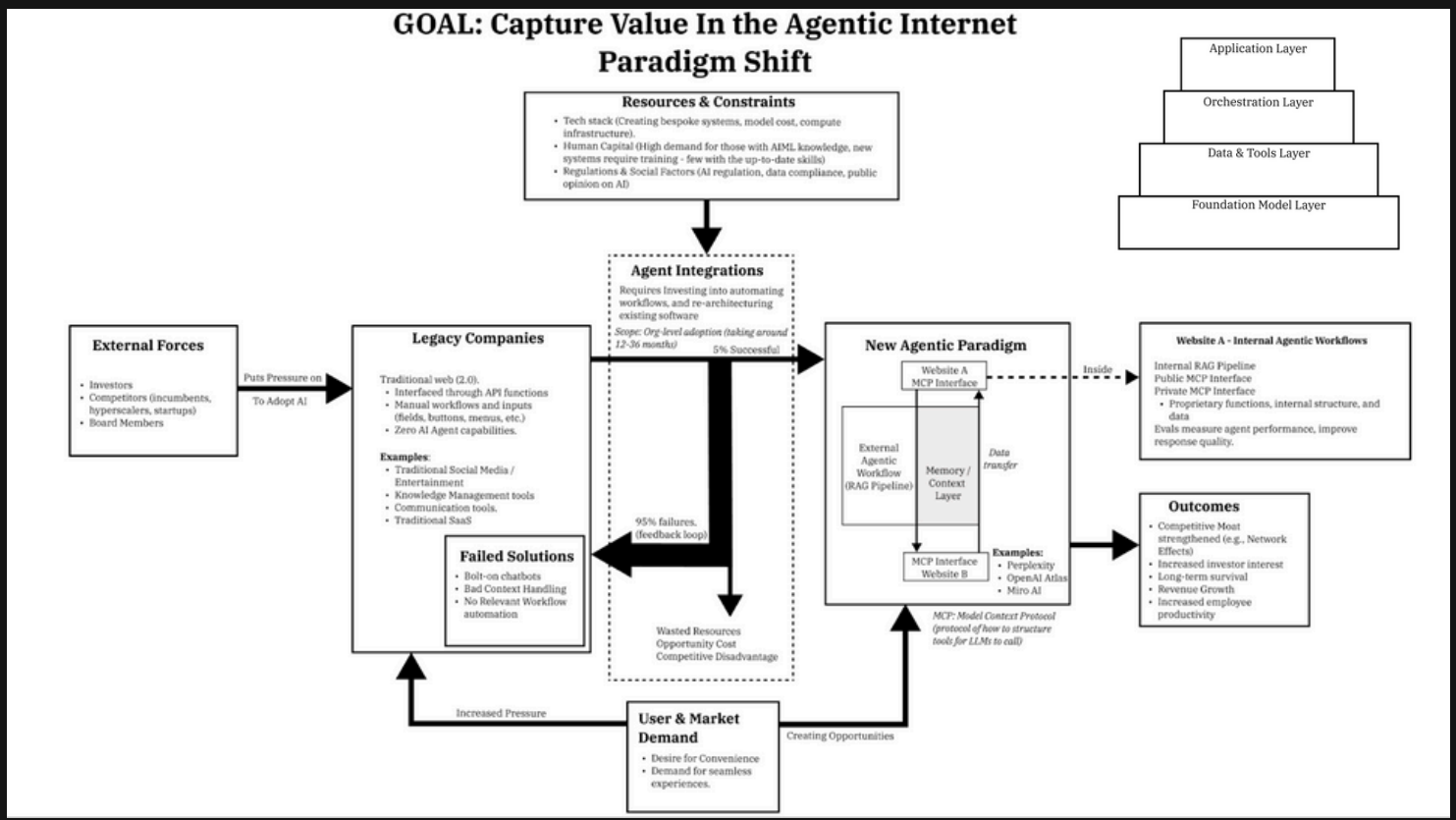


Figure 3: Systems map of the agentic transition of enterprise applications.

Example of raw data:

{"type":"excalidraw/clipboard","elements":[{"id":"ZlafmgLDuekfpjuriFz4x","type":"rectangle","x":677.5685356957075,"y":1049.964473057853,"width":177.66392890238615,"height":175.48403655537254,"angle":0,"strokeColor":"#1e1e1e","backgroundColor":"transparent","fillStyle":"solid","strokeWidth":2,"strokeStyle":"solid","roughness":1,"opacity":100,"groupIds":[],"frameId":null,"index":"ad","roundness":{"type":3},"seed":1608874298,"version":18,"versionNonce":809649382,"isDeleted":false,"boundElements":[],"updated":1761660477003,"link":null,"locked":false}],"files":{}}

This is for a single rectangle in Excalidraw, note how quickly the complexity would scale with larger systems

The Era of Agents

AI Agents are the future. The first Agent-first website and tools are being created: **Agentic Browsers** (Comet, Atlas, Dia), **Programming** (Codex, Cursor, Claude Code), **Search** (Perplexity, Exa), **Knowledge Work** (Notion, Microsoft Copilot), **Workflow Automation** (n8n, Zapier).

Self-Improving AI

Achieving Artificial General Intelligence, representing AI parity with humans. AI models emerge with unified Language, Spatial, Physical, Reasoning capabilities ('World Models').

Fully Agentic Web

Every website will have a LLM interface, communicating information with each other. An example of this is ChatGPT Agent mode, where it can order online on the user's behalf.

Dynamic Web

Frontend (visual styling) and Backend (Functionality) will be served by LLMs in real-time. This is "Code on demand" as mentioned by Sam Altman.

Causal Layered Analysis:

Litany: AI hallucinations, unreliable outputs, wasted budgets.
Systemic causes: Linear text-only architectures, lack of spatial encoding standards, RAG pipelines misaligned with graph data.
Worldview: Text = knowledge. Inherited from Natural Language Processing.
Myth/Metaphor: LLMs as blind librarians. Vast archives, but cannot perceive the visual layout.
Transformative future metaphor: LLMs as cartographers of meaning, seeing and reasoning across spatial knowledge.

AI capabilities are advancing faster than industries are adapting, resulting in increased opportunities for startups. This signifies a paradigm shift in how we interface, interact, and communicate across the internet. From user-based to chat box and agent-based. Legacy systems face being excluded by the agentic web if their data is not structured correctly for an LLM to search over — such companies could face their traffic effectively **dropping to zero** as they would be excluded from the agent network.

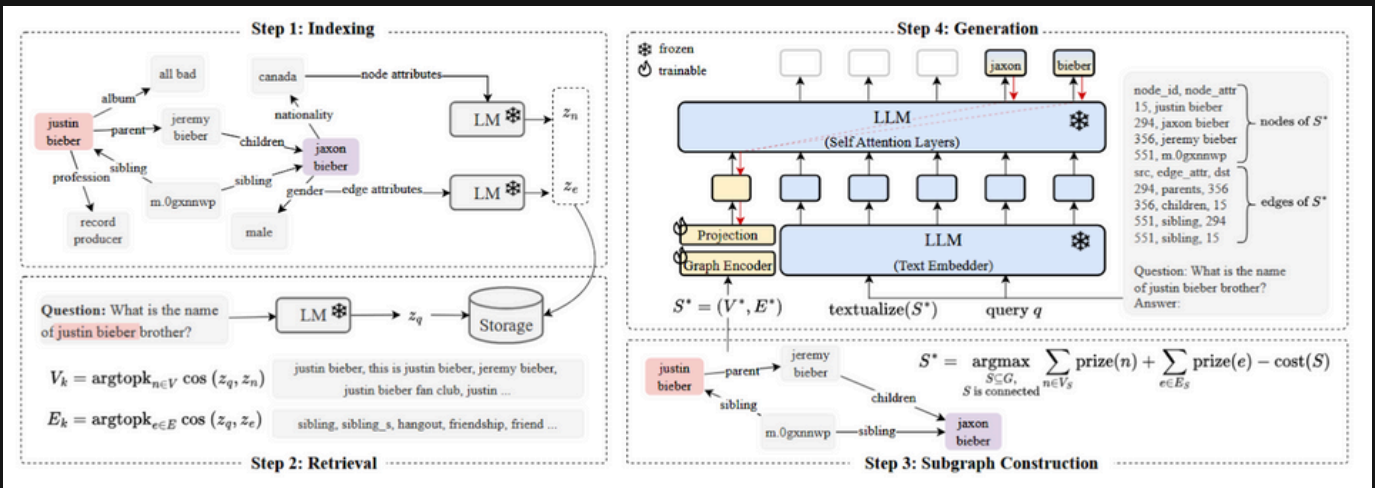


Figure 4: GraphRAG / G-Retriever Pipeline. (He et al., 2024).

Graph-oriented retrieval as a solution keeps intrinsic data relationships, but implementation remains complex for non-specialised teams.

The context component is a small part of the AI Agent system, but is critically important to the accuracy of results. We need to create a way to encode this information for LLMs (or more broadly, AI systems) to understand so the **latent meaning** of **spatial information** is not lost.

Enterprises *can't* do AI

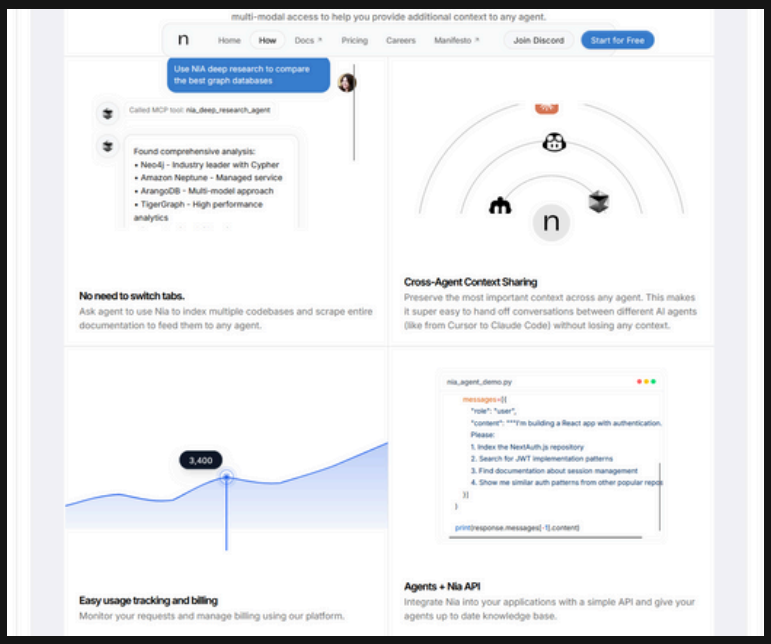
Despite estimated enterprise investment of **$30-40 billion** into Generative AI, businesses are failing to see meaningful returns. The agents are interacting with irrelevant data by missing the human aspect. **~95%** of organisations are getting zero measurable return (Challapally et al., 2025). Only **~5%** of integrated AI pilots are extracting meaningfully large value (millions of dollars).

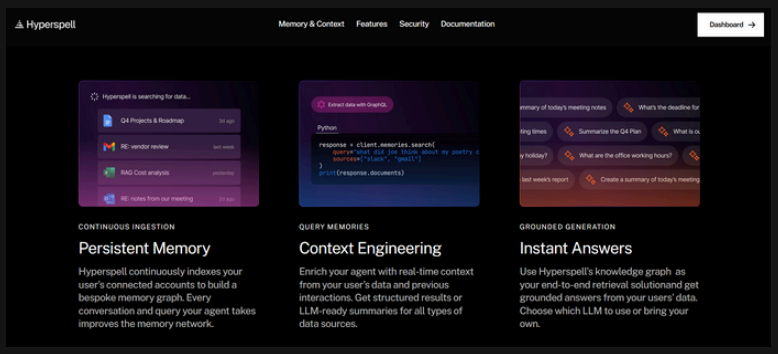
Miro, despite having **100M+ users** (Miro, n.d.) emphasises the text content of sticky notes when generating its responses, ignoring the other dimensions (colour, size, type, etc.) even when they are obvious to the observer. AI features include sticky-note summarisation and clustering. They have recently announced 'Visual Context Processing', highlighting the demand for structured intelligence by developing their own proprietary system (Miro.com, 2025).

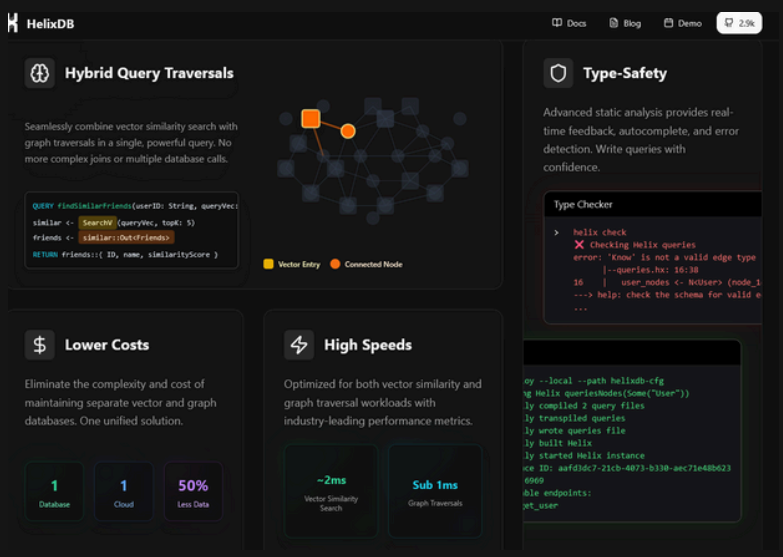
RAG is the dominant AI architecture approach, rising **from 31% to 51%** (Tully, 2024), but more than a third of surveyed organisations do not have a clear vision of how generative AI will be implemented across the enterprise (Tully, 2024). They are trying to add AI to their existing systems instead of redesigning their architecture to a more agent-friendly design. Data for LLMs is useless if it is **unstructured** and **self-contradictory.**

Because this problem space is emerging, there is currently limited empirical data quantifying the loss of spatial or relational information in enterprise RAG systems. This itself is a gap and opportunity.

An emerging approach is for startups to implement / provide infrastructure using low-latency graph-vector databases, providing improved context layers for LLMs (examples include **Nozomio** (https://www.trynia.ai), **Hyperspell** (https://www.hyperspell.com), **HelixDB** (https://www.helix-db.com))







All of these startups have been accepted by Y Combinator and show clear potential in reducing the LLM hallucination problem. They each solve different problems to the one I have identified, and if I am successful, they would adopt my solution.

What does a solution *look* like

Target Audience - Businesses / startups that want the following:

- Reduced hallucinations
- Higher retrieval accuracy
- Lower LLM / labour costs (token usage)
- Ability to continue using existing systems
- To leverage intrinsically structured data

The consumer layer is competitive (Miro, top 50 private company worth **$17.5bn** (CB insights, 2023), is adopting AI proactively into their canvas systems (Miro.com, 2025)). 'AI wrappers' are hard to stand out even with novel architecture due to a swathe of low effort products.

Will this be a problem in the future if LLMs keep improving? Foundational Research Labs (Deepmind, Anthropic, OpenAI) are always looking to solve these types of problems to improve their models. There is little public information on whether they are also working in the problem space identified. Or in the context layer. **Context is underserved** by the main research labs and instead they are focusing on consumer experiences & direct API solutions.

Multimodal models can process images but:

- Don't preserve underlying structured data relationships
- Expensive (vision tokens cost 10-20x text)
- Still require external retrieval systems
- The context layer remains infrastructure problem regardless of model improvements

A solution could be to build the tools and infrastructure specifically related to structured or spatial information. A feasible and common approach I would for a solution would be to focus on a dev-first, open-source SDK, following the likes of n8n, LangChain, mem0.

Why open-source SDK?

- **Network Effects**: Need widespread adoption for it to become the standard.
- **Enterprise Trust**: Adds transparency by keeping code public and accessible.
- **Adoption Rate**: LangChain reached 118k stars after three years of being open-source (Langchain, 2022).
- **Competitive Moat**: First-mover advantage being the open-source structural standard for RAG.
- Attracts talented volunteer contributors.

Why developer-first?

- Bottom-up adoption: with engineers and future founders, scaling later to enterprise.
- Degree of SDK lock-in (be the easiest to use and learn in the market).
- Easier to understand what developers want, as a software developer myself.

Why Me? As a developer with a background creating RAG systems and in Machine Learning, I've experienced this problem with LLMs. I understand the frustration of trying to make AI work with complex and interconnected data and am positioned to build a solution that developers will actually want to use.

Why Now? The rapid advance of agentic LLMs has created a critical infrastructure gap. The market is seeking a solution to the context problem today before billions more are wasted. Context is currently an area of high interest for investors (such as Y Combinator (Y Combinator, 2025))

What cannot be done? Training, data, and compute resources are all challenging (resource constraints and competition), so are outside my direct control (at least at this stage).

Within 5 years, every enterprise system will have an AI agent interface. These agents will need to navigate knowledge graphs, spatial workflows, and structured hierarchies.

The opportunity: Build the spatial encoding standard that becomes as fundamental to agentic AI as HTTP was to the web browser.

References

Anthropic.com. (2024). Building Effective AI Agents. [online] Available at: https://www.anthropic.com/engineering/building-effective-agents.

Anthropic.com. (2025). A small number of samples can poison LLMs of any size. [online] Available at: https://www.anthropic.com/research/small-samples-poison.

CB insights (2023). The Complete List of Unicorn Countries. [online] Meet all of the world's startups valued at $1B+ in our ongoing list of all unicorn companies in the U.S. and abroad. Available at: https://www.cbinsights.com/research-unicorn-companies.

Challapally, A., Pease, C., Raskar, R. and Chari, P. (2025). The GenAI Divide STATE OF AI IN BUSINESS 2025 MIT NANDA. [online] Available at: https://mlq.ai/media/quarterly_decks/v0.1_State_of_AI_in_Business_2025_Report.pdf.

Google Cloud (2024). What Is Retrieval Augmented Generation (RAG)? | Google Cloud. [online]

Google Cloud. Available at: https://cloud.google.com/use-cases/retrieval-augmented-generation. Google Cloud (2025). What are AI agents? Definition, examples, and types. [online] Google Cloud. Available at: https://cloud.google.com/discover/what-are-ai-agents.

He, X., Tian, Y., Sun, Y., Chawla, N., Laurent, T., Lecun, Y., Bresson, X. and Hooi, B. (2024). G-Retriever: Retrieval-Augmented Generation for Textual Graph Understanding and Question Answering. [online] Available at: https://proceedings.neurips.cc/paper_files/paper/2024/file/efaf1c9726648c8ba363a5c927440529-Paper-Conference.pdf [Accessed 30 Jun. 2025].

Kwa, T., West, B., Becker, J., Deng, A., Garcia, K., Hasin, M., Jawhar, S., Kinniment, M., Rush, N., Arx, S., Bloom, R., Broadley, T., Du, H., Goodrich, B., Jurkovic, N., Miles, L., Nix, S., Lin, T., Painter, C. and Parikh, N. (2025). Measuring AI Ability to Complete Long Tasks. [online] Available at: https://arxiv.org/pdf/2503.14499

Langchain (2022). LangChain. [online] GitHub. Available at: https://github.com/langchain-ai/langchain.

Miro (n.d.). About Miro | Meet the team | Our mission. [online] https://miro.com/. Available at: https://miro.com/about/

Miro.com. (2025). Miro AI overview . [online] Available at: https://help.miro.com/hc/en-us/articles/28765406244498-Miro-AI-overview.

OpenAI (2025). Announcing The Stargate Project. [online] Openai.com. Available at: https://openai.com/index/announcing-the-stargate-project/

Tully, T. (2024). 2024: The State of Generative AI in the Enterprise - Menlo Ventures. [online] Menlo Ventures. Available at: https://menlovc.com/2024-the-state-of-generative-ai-in-the-enterprise/

Y Combinator (2025). The YC Startup Directory | Y Combinator. [online] Y Combinator. Available at: https://www.ycombinator.com/companies?batch=Fall%202025&batch=Summer%202025&industry=Engineering%2C%20Product%20and%20Design [Accessed 28 Oct. 2025].