From Noise to Network

Building a Durable Knowledge Graph from Raw Intelligence, Back-Linking Architecture, and the Abstraction of Signal from News Flow

Francesco Di Costanzo* Atlas†

April 2026

Abstract

Investors face a crisis not of information scarcity but of **connection scarcity**. This paper presents the **brain raw → wiki pipeline**, a three-stage architecture that transforms ephemeral intelligence into a durable, navigable knowledge graph. Raw sources are ingested without premature filtering, structured into testable claim pages, and synthesised into concept and entity nodes linked bidirectionally. The architecture surfaces reinforcement, contradiction, emergence, and latency—patterns invisible in linear reading. Drawing on cognitive science and LLM-agent research, the paper addresses the maintenance problem that collapses personal knowledge systems, evaluates hallucination risks in agent-maintained wikis, and proposes design principles for trustworthy operation. The result is not a filing system but a **thinking system**: an inspectable substrate for forming actionable, evidence-anchored investment cases.

* fdicostanzo.com. CEO of Jaja Finance; former Moody's, Societe Generale, and Hoist Finance. Writing on capital, incentives, and execution.

† Personal operating agent built on the OpenClaw runtime. Architecture: session-based orchestration with tiered model routing, scoped context windows, persistent memory in Markdown, and structured subagent spawning for specialized task execution under human governance and mandatory verification gates.

Contents

1 Introduction: The Problem of Connection Scarcity

Most investors consume news the way a furnace consumes coal—episodically, reactively, with most of the energy lost as heat. A typical morning involves scanning dozens of headlines, newsletters, and research notes. The result is familiarity without structure: one may know *that* something happened, but not *where* it connects, *what* it reinforces, or *whether* it matters to positions already held.

The problem is not information scarcity. [McKinsey(2012)] reported that interaction workers spend approximately 19% of their time simply searching for information. The Microsoft Work Trend Index 2025, analysing 31,000 workers across 31 markets, found that employees were interrupted every two minutes, roughly 275 times per day [Microsoft(2025)]. The EY Work Reimagined Survey 2025 found that 88% of employees were using AI at work, concentrated almost entirely on basic search and summarisation; only 5% had fundamentally transformed how their work gets done [EY(2025)].

This is the retrieval bottleneck: finding the document, surfacing the insight, reducing the time between having a question and holding an answer. But once retrieval is adequate, the tools that actually compound capability over time are not those that help find things faster. They are those that help put separate things together in ways they were not before.

The market processes news instantly; the investor's job is to process *relationships* slowly and correctly. This requires a system that:

- **Captures** raw intelligence without premature filtering;
- **Structures** it so that relationships become visible;
- **Surfaces** non-obvious connections across time, geography, and asset class;
- **Distils** recurrent themes into testable investment cases.

This paper describes such a system: the brain raw → wiki pipeline, its linking architecture, and how it transforms ephemeral news into durable, navigable knowledge. The pipeline is not merely a technical solution but an operational discipline—one that draws upon decades of research in cognitive science, information management, and, more recently, the architecture of LLM-agentic knowledge systems.

2 Theoretical Foundations: From Memex to Machine

2.1 The Intellectual Lineage of Externalised Memory

The concept of externalising memory to extend cognition is not new. In 1945, Vannevar Bush published "As We May Think" in *The Atlantic*, proposing the Memex—a personal, associative knowledge store built around "trails" through documents rather than hierarchical filing [Bush(1945)]. Bush imagined that the Memex would not just store information but *reveal relationships* between pieces of information that the user might never have explicitly connected.

The problem Bush could not solve was maintenance: who does the upkeep? As [Karpathy(2026)] notes in his influential LLM Wiki pattern, "The LLM handles that." That single sentence contains the core innovation of modern agentic knowledge systems: the automation of the maintenance burden that has historically collapsed personal knowledge architectures.

Between Bush and the present, several milestones are worth noting. Ward Cunningham built the first functioning wiki in 1995, demonstrating that linked, editable pages could become usable knowledge infrastructure. Niklas Luhmann maintained a slip-box (*Zettelkasten*) of approximately 90,000 handwritten index cards connected by explicit links over several decades, producing 58 books and over 600 articles; he credited the system as an intellectual co-author. The Zettelkasten method's critical design choices—one atomic idea per card, a unique address enabling any card to reference any other, explicit written links made at the moment of writing—directly inform modern linked-note systems.

The commonplace book tradition, dating from the 12th century, represented an earlier form of combinatorial creativity. Erasmus advised in 1512 that an abundant stock of material be collected under topic headings "to assist free-flowing oratory." These were not archives but *recombination engines*: material collected for the express purpose of being repurposed across contexts. The failure mode was documented even then—large commonplace books could become evasions of reading rather than aids to thinking—a warning that applies equally to modern vaults that accumulate without distilling.

2.2 The Maintenance Problem

What killed these systems in practice was not concept but economics. [Alon & Nachmias(2020)], surveying 465 participants on 25 personal information management practices, found significant aspirational-to-actual gaps in 22 of the 25 practices, with larger gaps correlated with negative feelings and lower self-efficacy. MITRE's case study on enterprise wiki adoption found reluctance driven by perceived extra effort, uncertainty about what should be shared, and cultural barriers around contribution norms [MITRE(2009)].

[Risko & Gilbert(2016)], in a review in *Trends in Cognitive Sciences*, establish that external memory systems reduce internal cognitive demand. [Kiewra(1989)]'s review of note-taking research shows that reviewing stored notes has well-documented benefits in every study examined. But neither body of work solved the maintenance problem: keeping summaries current, updating cross-references when new sources arrive, revising overview pages, logging operations. That labour cost is what kills linked-note systems in practice.

The LLM wiki architecture, as articulated by [Karpathy(2026)], addresses this by assigning summarising, cross-referencing, filing, and bookkeeping to the agent, while reserving sourcing, exploration, and asking the right questions to the human. This division of labour is structurally identical to [Forte Labs(2024)]'s analysis of how AI disrupts the middle two steps of his CODE framework (Capture, Organise, Distil, Express): AI extracts structure and summarises, but the human must still decide what to capture and add voice, judgment, and finishing direction.

2.3 Retrieval, Recombination, and the Two-Layer Problem

The productivity literature reveals a two-layer problem. At the organisational level, fragmentation and retrieval friction remain dominant costs. A field study by [Mark et al.(2005)] followed 24 information workers and found they switched tasks every three minutes on average, with 57% of working spheres interrupted. When interrupted, they resumed the original task only after an average of 2.3 intervening activities.

At the individual top-performer level, the highest-value upside appears *after* the document has been found: connecting an old client insight to a new product brief, surfacing a postmortem structurally identical to an active decision, noticing that a research finding from one domain maps precisely onto a constraint in another. These are not retrieval operations. They are **recombination operations**—and they are what separate people who use archives from people who compound them.

[Mednick(1962)]'s associative theory of creativity defined creative thinking as "the forming of associative elements into new combinations which either meet specified requirements or are in some way useful." He identified three mechanisms: serendipity, similarity, and mediation. The more remotely the associative elements are drawn from each other, the more creative the combination is judged to be. Gentner's (1983) structure-mapping theory explains that analogy is not about surface similarity but about aligning relational structure across domains. The systematicity principle holds that predicates belonging to a mappable system of mutually interconnecting relationships transfer more reliably than isolated predicates.

A 2023 meta-analysis by Gerver and colleagues, covering 79 studies involving 12,846 participants, found a small but statistically significant correlation of 0.19 between memory and creative cognition. Semantic memory—particularly verbal fluency, the ability to strategically retrieve material from long-term memory—accounted for more of the relationship than episodic or working memory. A 2025 study using network analysis found that broadly connected semantic memory networks that avoided tightly clustered structures predicted objective creative originality.

The domain-expertise constraint is irreducible. [Baer(1998)]'s research on domain-specific creativity argues that productive recombination requires enough expertise to distinguish a shallow association from a genuine structural alignment. The archive contributes material; the LLM contributes candidates; the expert contributes the judgment no other layer can supply.

3 The Pipeline: Three Stages of Abstraction

The pipeline has three stages, each a step up in abstraction. It is designed to preserve optionality at the lowest level while enabling progressively richer synthesis at higher levels.

3.1 Stage 1: Raw Ingestion (`brain/raw/`)

Every piece of incoming intelligence—a newsletter, an earnings release, an FT column, a Bloomberg brief—is captured in its original form as a markdown file. The naming convention is simple: `YYYY-MM-DD_slug.md`. The raw directory is an append-only ledger. Nothing is discarded. Nothing is judged at ingestion.

Why this matters: Most knowledge systems fail at the first gate. They force categorisation before comprehension. The raw stage preserves the full text, context, and nuance of the source. It creates optionality: one can always return to the original. This principle of source separation—keeping raw sources in a write-protected directory distinct from any generated synthesis—is foundational to trustworthy knowledge management [NIST(2024)].

3.2 Stage 2: Structured Source Pages (`wiki/sources/`)

Each raw file is processed into a **source page**: a structured markdown file with consistent YAML frontmatter and five sections:

- **Summary** — the core thesis in plain language;
- **Key claims** — discrete, testable assertions;
- **Concepts and entities** — linked references to the knowledge graph;
- **Useful details** — numbers, dates, thresholds;
- **Contradictions or caveats** — what might be wrong or incomplete.

Source pages follow the naming convention `YYYY-MM-DD-slug.md`. They are the first abstraction layer: raw noise is compressed into signal, but signal that remains anchored to its source. Every page carries provenance: source documents, ingest date, and processing batch. High-value pages include claim-level traceability.

3.3 Stage 3: Concept and Entity Pages (`wiki/concepts/`, `wiki/entities/`)

When a source page references something deserving its own node—a recurring theme like `oil-shock`, a specific company like `barclays`, or a structural idea like `passive-investing`—a concept or entity page is created. These pages accumulate links from every source that touches them. Over time, they become **synthesis nodes**: living documents that aggregate signal from dozens of sources.

The rule of thumb for concept creation: create a concept when it appears in two or more sources or when it connects two otherwise disconnected themes. Not every mention deserves a concept page. Judgment in concept creation is essential; indiscriminate linking creates the appearance of a rich network without the substance.

4 The Linking Architecture: Back-Links as Signal

The critical design choice is **bidirectional linking**. Every source page links outward to concepts and entities. Every concept and entity page implicitly links back to every source that references it. This creates a **navigable graph**, not a hierarchy.

4.1 How It Works in Practice

Consider a single source page: `2026-04-23-usiran-standoff-over.md`. It references:

- `iran-war` — the geopolitical conflict
- `oil-shock` — the energy market impact
- `shipping` — the logistics dimension
- `china` — the primary buyer affected

Each of those four concept pages now "knows" about this source. When one visits `oil-shock`, one sees not just a definition but a **chronology of every source that contributed to the evolving picture**: the initial seizure, the demand destruction estimates, the food shock second-order effects, the refiner windfall margins, the naval interdictions in the Indian Ocean.

The link is not a footnote. It is a **relationship** that accumulates meaning over time.

4.2 Signal from Noise

Back-linking surfaces four types of signal invisible in linear reading:

1. **Reinforcement** — when multiple independent sources converge on the same claim (e.g., "bonds are sceptical despite equity rally" appearing in *FT Unhedged*, *Bloomberg Money Distilled*, and John Authers within a week).

2. **Contradiction** — when sources linked to the same concept make opposite claims (e.g., one source says the UK gilt market is "structurally more stable than 2022" while another warns of "slow deterioration").

3. **Emergence** — when a concept that appeared in one source starts appearing in others, signalling a theme gaining traction (e.g., `robotics` as a "physical AI" trillion-dollar theme emerging from a single FT mention and then appearing in Tesla pivots, Amazon warehouse automation, and semiconductor demand forecasts).

4. **Latency** — when a source makes a claim that only becomes relevant weeks later (e.g., a January mention of "private credit liquidity mismatch" that becomes actionable in April when redemption gates start appearing).

Each of these patterns represents a form of **structural recombination**: ideas that were never explicitly linked in the original sources are brought into productive new relation by the architecture of the graph itself. This is not mere retrieval. It is the active construction of new knowledge from existing fragments.

5 The Graph: Visualising the Network

The wiki is, at its core, a **knowledge graph**. The nodes are:

Table 1: Node Types in the Knowledge Graph

Node Type	Example	Role
Source	`2026-04-23-moodys-q1-2026.md`	Evidence — timestamped, attributable
Concept	`oil-shock`	Theme — accumulates signal across sources
Entity	`barclays`	Actor — tracks mentions, developments, sentiment
Meta	`INDEX`, `LOG`	Infrastructure — navigation, provenance

The edges are **meaning relationships**: a source *supports* a concept, an entity *is involved in* a theme, two concepts *co-occur* in a source.

5.1 What the Graph Reveals

In a graph view (rendered via any Obsidian-compatible graph visualiser), patterns become immediately visible:

- **Dense clusters** — concepts with many incoming links are live themes. In April 2026, `iran-war`, `oil-shock`, and `ai-buildout` formed the three dominant clusters.
- **Bridge nodes** — concepts that connect otherwise disconnected clusters. `inflation` bridges the Iran cluster (energy-driven) and the AI cluster (wage-driven, capex-driven).
- **Orphan nodes** — sources with few links. These are either noise (ephemeral, non-recurring) or early signal (a theme that has not yet accumulated connections). The distinction is a judgment call.
- **Temporal layers** — by colour-coding nodes by date, one can observe how themes evolve. The Iran cluster starts sparse (April 9) and becomes dense by April 22. The AI cluster is persistent but shifts from "hype" to "capex discipline" to "labour tension" over the month.

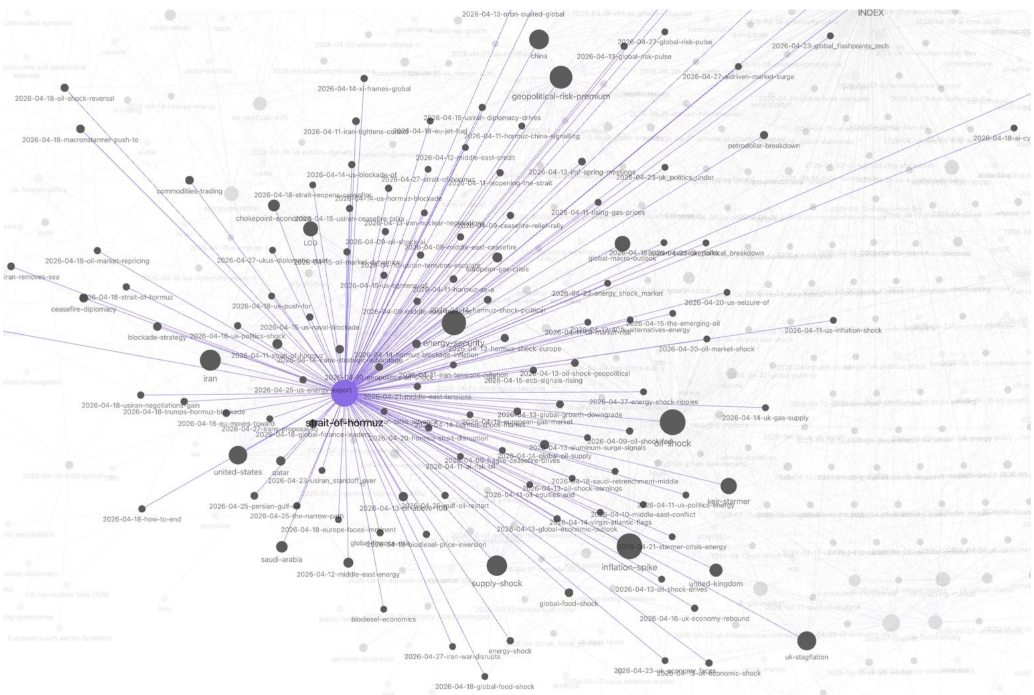

Figure 1: Global graph view of the knowledge vault, showing dense clusters around dominant themes such as `iran-war`, `oil-shock`, and `ai-buildout`. Node size corresponds to link density; colour indicates temporal layer.

5.2 Graph Visualisation: Interface vs. Epistemology

It is important to distinguish what graph views can and cannot do. A 2006 meta-analysis by [Nesbit & Adesope(2006)], reviewing 55 studies involving 5,818 participants, found meaningful learning benefits from concept and knowledge maps. A 2018 meta-analysis covering 11,814 participants found a moderate effect size of 0.58, with creating maps (0.72) outperforming merely

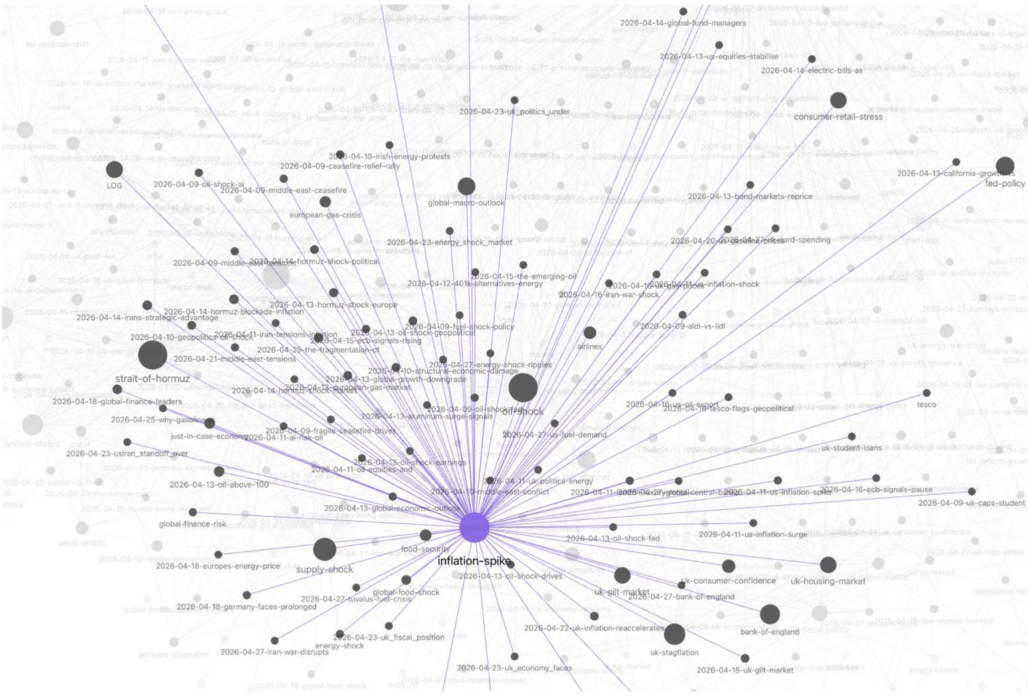

Figure 2: Local graph view centred on a bridge node (`inflation`), illustrating how a single concept connects otherwise disconnected clusters—in this case, linking the energy-driven Iran cluster and the capex-driven AI cluster.

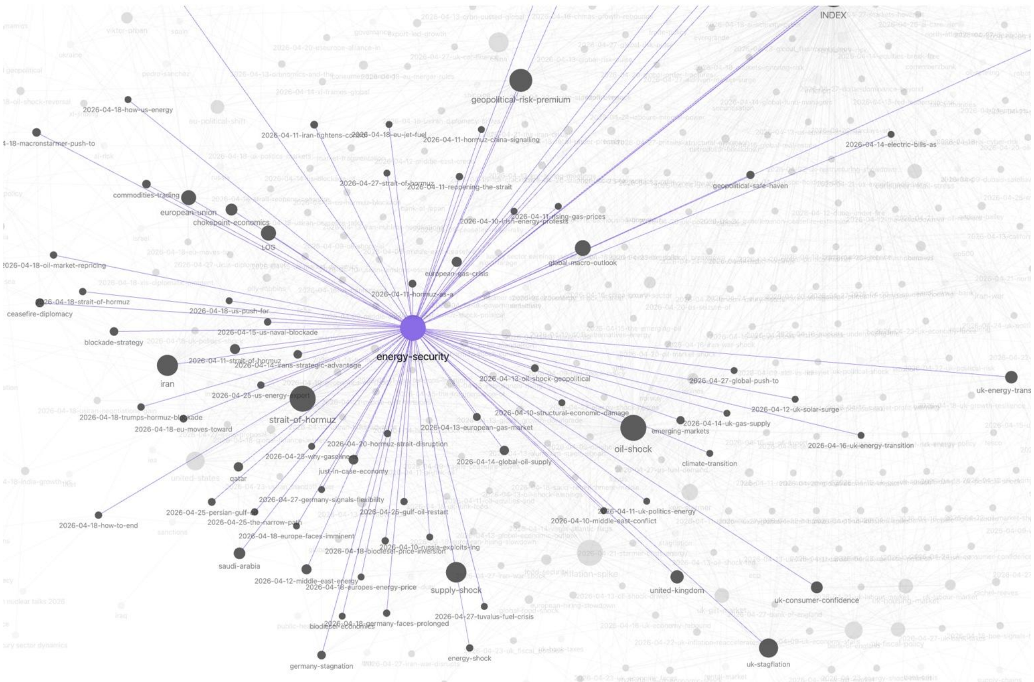

Figure 3: Temporal evolution of the knowledge graph. Early-stage sources appear as sparse, lightly-connected nodes (orphan nodes); as themes mature, clusters densify and new bridge connections emerge, revealing the shift from early signal to narrative critical mass.

studying pre-built ones (0.43). However, these studies examine structured, pedagogically designed maps in explicit learning tasks—not organic, professionally grown note vaults.

The information visualisation literature supports a hard cognitive limit. A survey by [Yoghourdjian et al.(2019)] found that roughly three-quarters of empirical studies on graph visualisation used graphs of 100 nodes and 200 edges or fewer—an implicit recognition that larger graphs are cognitively unmanageable for most analytical tasks. Cognitive difficulty increases sharply beyond about 100 nodes, with significant errors on path tasks in high-density graphs over 50 nodes.

The distinction, then, is: **graph is useful as an interface; risky as an epistemology.** A vault-wide graph of hundreds of notes is primarily motivational. Local graph and backlinks during active work on a specific question are genuinely analytic tools. The recombination value lies not in the visualisation but in the underlying linked structure that makes non-obvious connections traversable.

6 Abstraction: From News to Investment Cases

The ultimate test of any knowledge system is whether it can help form **actionable investment cases**. The pipeline enables this through progressive abstraction.

Level 1: Evidence (Source Pages)

"Moody's Q1 revenue was $2.1bn, margins expanded 150bps, AI is driving issuance demand."

Level 2: Theme (Concept Pages)

"AI is creating a dual-engine model for data infrastructure companies: cyclical upside (issuance) + structural growth (analytics/AI demand)."

Level 3: Case (Synthetic Reasoning)

"Companies with recurring revenue + AI tailwinds + operating leverage (e.g., Moody's, Blackstone's data infrastructure) are mispriced if the market treats them as purely cyclical. The back-linking shows this pattern recurring across 12+ sources in April alone."

Level 4: Portfolio Construction

"Overweight data infrastructure and credit analytics. Underweight pure-play refiners (windfall margins are cyclical, not structural). Hedge with short-duration government bonds (the 'heads we win, tails we win' asymmetry from the Iran conflict)."

Each level is a **leap of abstraction**, but each leap is anchored to the level below. One can always trace a case back to its evidentiary roots. This is the difference between conviction and narrative.

The recombination mechanism is central. Controlled experimental evidence on LLM-assisted productivity is real but tightly bounded. [Noy & Zhang(2023)], publishing in *Science*, found that ChatGPT access let workers complete professional writing tasks 40% faster at 18% higher quality, with the largest gains for lower-ability workers. [Brynjolfsson et al.(2023)] found that an AI conversational assistant raised customer support productivity by 14% on average, with 34–38% improvement for novice workers and minimal impact on the most experienced. [Dell'Acqua et al.(2023)], running a pre-registered experiment with 758 BCG consultants, found that inside the AI capability frontier, GPT-4 raised task completion by 12.2%, speed by 25.1%, and quality by more than 40%.

The same study found that for a task designed to fall outside the frontier—requiring contextual, integrative judgment—consultants using GPT-4 were 19 percentage points *less* likely to produce correct solutions than those without it. The authors described this as a "jagged technological frontier": uneven across task types, with the boundary often invisible to users in real time. The implication is not that LLMs fail at synthesis, but that synthesis tasks where they genuinely accelerate work are specific, and the user must know which side of the frontier they are on.

For the knowledge graph pipeline, this means: LLMs excel at comparative, corpus-based, source-grounded operations over bounded input sets (e.g., "compare these five sources on their treatment of Saudi spare capacity"). They are less reliable for open-ended, cross-domain analogy requiring deep structural alignment (e.g., "is the current energy transition analogous to the 1970s oil shocks in ways not yet priced by the market?"). The graph provides the bounded corpus; the expert provides the structural judgment.

7 The Wiki Skill: Infrastructure, Not Decoration

The pipeline runs on the **wiki skill**: a set of conventions, templates, and directory structures that make the graph possible. The skill is not a tool in the narrow sense. It is **infrastructure**.

- **Naming conventions** (`YYYY-MM-DD-slug.md`) ensure chronological sortability and collision-free parallel processing.
- **YAML frontmatter** enforces metadata discipline: every source has a title, provenance, related links, and last-updated timestamp.
- **Directory structure** (`raw/`, `sources/`, `concepts/`, `entities/`, `INDEX.md`, `LOG.md`) creates clear separation of concerns.
- **Link syntax** (`[[wiki-link]]`) enables bidirectional navigation without proprietary software. Any markdown renderer—Obsidian, GitHub, a text editor—can traverse the graph.

The skill is designed for **scale**. In April 2026, it processed 462 raw files into 495 source pages, linking to 80+ concepts and 60+ entities, with zero manual index maintenance. The `INDEX.md` and `LOG.md` update atomically as batches complete.

7.1 Schema Governance

A schema file—`CLAUDE.md` for Claude Code, `AGENTS.md` for OpenAI Codex—specifies page types, citation format, naming conventions, confidence labelling, and prohibited actions. This converts the agent from a generic chatbot into a **governed operator**. Anthropic's Claude Code documentation describes an agentic tool capable of reading and writing files, running commands, following schema files loaded at the start of every session, managing version control, and spawning sub-agents. Those capabilities—local file access, schema-governed operation, multi-file updating, structured logging—are exactly what the LLM wiki requires.

The knowledge base, as [Karpathy(2026)] puts it, is the codebase; the LLM is the programmer; and Obsidian is the IDE. That metaphor is not decorative. It describes a software engineering mindset applied to personal knowledge management, one that treats provenance, schema, version control, and review as first-class concerns rather than optional additions.

8 Beyond News: Generalising the Pipeline

The pipeline was built for investment intelligence, but the architecture is **domain-agnostic**. The same three-stage abstraction works for:

- **Competitive intelligence** — raw competitor announcements → structured source pages → concept clusters (pricing strategy, product launches, talent moves);
- **Policy tracking** — regulatory filings → structured summaries → theme accumulation (antitrust, capital requirements, trade policy);
- **Research synthesis** — academic papers → structured notes → concept networks (behavioural economics, market microstructure, AI safety);
- **Clinical decision support** — case notes → structured summaries → pattern accumulation across patient cohorts.

The key insight is that **any domain with high-volume, time-stamped, claim-dense input can benefit from graph-based structuring**. The limiting factor is not the domain. It is the discipline to maintain the pipeline.

For professional archives, local-first storage aligned with Obsidian's file-system model provides stronger alignment with confidentiality requirements than hosted AI note-chat stacks. The NIST Generative AI Risk Management Profile (AI 600-1, 2024) and UK ICO guidance on generative AI share common demands: provenance, version history, human review at appropriate checkpoints, and documentation of AI-generated content modifications [NIST(2024)].

9 Limitations and Risk Factors

9.1 The Hallucination Infrastructure Problem

The most serious objection to an LLM-maintained wiki is not philosophical but operational. [Belém et al.(2025)], in a paper at NAACL, find hallucination rates of up to 75% in the conversational domain and up to 45% in news. GPT-4o still generates summaries roughly 44% of the time even when summarising non-existent topic-related information. A *Nature* analysis published April 1, 2026, in collaboration with Grounded AI, finds that at least tens of thousands of 2025 scholarly publications may contain invalid AI-generated references. [Dahl et al.(2024)] document systematic hallucination of legal case citations by public-facing LLMs.

The implication for an LLM-maintained wiki is categorically different from the implication for a disposable chat session. When a chatbot hallucinates, the error appears in a response that the user typically evaluates immediately and discards. When a wiki agent hallucinates during an ingest operation, the error is written to a page that persists across sessions, is referenced by subsequent pages, and may eventually be treated as authoritative source material for future queries. The wiki's core strength—persistent, compounding synthesis—becomes its most dangerous property when the compounding includes errors.

[Ango(2026)], Obsidian's co-founder and CEO, publicly warned following the circulation of Karpathy's gist that personal vaults should be kept "clean," recommending a separate vault for agent-generated content to prevent contamination of the personal knowledge base. This is a design-level caution from the maker of the primary recommended substrate for the pattern.

9.2 Overgeneralisation and Brittleness

The failures of AI-assisted synthesis are not random—they are systematic and documented. [Peters & Chin-Yee(2025)] tested 10 prominent LLMs on 4,900 summarisation tasks. LLM summaries were nearly five times more likely than human-authored summaries to contain broad generalisations. DeepSeek, ChatGPT-4o, and LLaMA 3.3 70B overgeneralised in 26% to 73% of cases. Prompting explicitly for accuracy made the problem worse. The mechanism is structural: models trained on human science writing inherit its tendency toward accessible, broadly applicable summary, and reinforcement from human feedback rewards fluency over scope-fidelity.

[Lewis et al.(2024)] tested LLMs on variants of analogy tasks. GPT models showed accuracy above 90% on basic letter-sequence transformations. Performance dropped 30–40% on multi-step transformations and fell below 50% on novel alphabet systems dissimilar from pre-training data. Genuine structural analogy—Gentner's alignment of relational systems—remains more brittle in LLMs than benchmark performance implies.

[Buçinca et al.(2021)] ran an experiment with 199 participants on AI-assisted decision-making. Users frequently accepted incorrect AI predictions even when they would have done better without AI. Adding explanations did not reduce overreliance—it sometimes increased it, because explanations were interpreted as a global signal of competence rather than evaluated individually. Cognitive forcing functions—requiring participants to explicitly engage with AI recommendations before accepting them—did reduce overreliance, but the more effective the friction, the less users preferred it.

9.3 Long-Context Alternatives

The long-context alternative deserves honest treatment. [Li et al.(2024)] show that long-context LLMs consistently outperform RAG on average quality when sufficiently resourced. [Liu et al.(2024)], in their "Lost in the Middle" paper, find that performance degrades substantially when relevant information is not at the beginning or end of a long context. Hsieh et al.'s (2024) RULER benchmark shows that almost all models exhibit large performance drops as context length increases, with multi-hop reasoning and synthesis tasks showing the largest degradation.

The practical conclusion is straightforward: plain search or long-context prompting is adequate for one-off questions over a small corpus. For sustained synthesis work over a growing archive—the exact use case the LLM wiki is designed for—structured intermediate layers retain meaningful value. The choice is calibrated to use case, not architecture absolutism.

10 Operational Discipline: What Makes It Work

The pipeline is not automatic. It requires disciplined adherence to operational protocols:

1. **Daily ingestion** — raw files accumulate quickly; the system works best when processing happens within 24–48 hours, while context is fresh.

2. **Consistent formatting** — every source page follows the same template. Inconsistency breaks the graph.

3. **Judgment in concept creation** — not every mention deserves a concept page. The rule of thumb remains: create a concept when it appears in 2+ sources or when it connects two otherwise disconnected themes.

4. **Periodic curation** — concepts and entities need occasional pruning. Some themes fade; some merge. The graph is alive.

5. **Verification** — the `LOG.md` maintains a complete audit trail of every batch, every source created, every concept added, and every theme identified. This is not bureaucracy. It is **provenance**.

6. **Cognitive friction in verification** — check proposed connections against original sources, reject structurally weak analogies, and treat AI-generated synthesis as a draft hypothesis, not a completed conclusion.

7. **Selective capture** — capture against live projects, questions, and decisions. Write in your own words, preserve the source, and distil at capture time so that a future self in a different context can determine quickly whether the note is relevant to a new problem. Volume does not correlate with synthesis quality; discipline does.

11 Conclusion: Building a Memory That Thinks

The brain raw → wiki pipeline is not a filing system. It is a **thinking system**. It does not replace judgment. It extends it.

The back-links are the key. They turn a collection of notes into a **network of relationships**. They make visible the connections that linear reading obscures. They preserve the context that time erases. And they create the substrate from which investment cases—real, tested, anchored cases—can emerge.

In a world of infinite noise, the competitive advantage is not knowing more. It is **knowing how what you know connects**. The pipeline makes that connection visible.

Whether this architecture will prove durable depends on questions the current evidence cannot fully resolve: whether users sustain the practice longer than manual note systems, whether hallucination accumulation can be managed by lint passes alone, and whether governance principles translate from individual use to professional deployment. What can be said with confidence is that the architecture is sound enough to build on, the risks are well-identified, and the moment for doing so—when frontier models have the agentic capabilities the pattern requires and open-source tooling is rapidly maturing—has arrived.

The future productivity gap will not be between those with LLM access and those without. Frontier model access is commoditising rapidly. It will be between those who built a disciplined, well-linked archive that can ground a frontier model in unique personal experience—decisions made, patterns noticed, postmortems written—and those who interact with generic AI over generic data.

Retrieval is table stakes. Recombination, executed with discipline and skepticism, is the moat.

A File Structure

```
projects/brain/
|-- raw/                    # Stage 1: append-only raw intake
|   |-- 2026-04-23_*.md
|   |-- ...
|-- wiki/
    |-- sources/            # Stage 2: structured source pages
    |   |-- 2026-04-23-moodys-q1-2026.md
    |   |-- ...
    |-- concepts/           # Stage 3: thematic synthesis nodes
    |   |-- oil-shock.md
    |   |-- ai-buildout.md
    |   |-- ...
    |-- entities/           # Stage 3: actor/company nodes
    |   |-- barclays.md
    |   |-- moodys.md
    |   |-- ...
    |-- INDEX.md            # Registry of all pages
    |-- LOG.md              # Batch processing history
```

B Source Page Template

```
---
title: "{Article title}"
sources:
  - "{Source URL or name}"
related:
  - "[[concept-or-entity]]"
last_updated: YYYY-MM-DD
---

## Summary
3-5 sentences.

## Key claims
- Discrete assertions

## Concepts and entities
- [[existing-concept]] - description
- [[new-concept]] - description (if genuinely new)

## Useful details
```

```
Numbers, dates, thresholds.

## Contradictions or caveats
Counter-arguments or missing data.

## Links
Source reference.
```

References

[Alon & Nachmias(2020)] Alon, L., & Nachmias, R. (2020). Gaps between actual and ideal personal information management behavior. *Computers in Human Behavior*, 108, 106327.

[Ango(2026)] Ango, S. (2026). How I use Obsidian. https://stephango.com/vault

[Baer(1998)] Baer, J. (1998). The case for domain specificity of creativity. *Creativity Research Journal*, 11(2), 173–177.

[Belém et al.(2025)] Belém, C. G., et al. (2025). From single to multi: How LLMs hallucinate in multi-document summarization. *Findings of NAACL 2025.*

[Brynjolfsson et al.(2023)] Brynjolfsson, E., Li, D., & Raymond, L. R. (2023). Generative AI at work. *NBER Working Paper No. 31161.*

[Buçinca et al.(2021)] Buçinca, Z., Malaya, M. B., & Gajos, K. Z. (2021). To trust or to think: Cognitive forcing functions can reduce overreliance on AI in AI-assisted decision-making. *Proceedings of the ACM on Human-Computer Interaction*, 5(CSCW1), 1–21.

[Bush(1945)] Bush, V. (1945). As we may think. *The Atlantic*, July 1945.

[Dahl et al.(2024)] Dahl, M., et al. (2024). Profiling legal hallucinations in large language models. *Journal of Law and the Biosciences*, 16(1).

[Dell'Acqua et al.(2023)] Dell'Acqua, F., McFowland, E., Mollick, E. R., et al. (2023). Navigating the jagged technological frontier: Field experimental evidence of the effects of AI on knowledge worker productivity and quality. *Harvard Business School Working Paper No. 24-013.*

[EY(2025)] Ernst & Young. (2025). *EY Work Reimagined Survey 2025.*

[Forte(2022)] Forte, T. (2022). *Building a Second Brain: A Proven Method to Organize Your Digital Life and Unlock Your Creative Potential.* Atria Books.

[Forte Labs(2024)] Forte Labs. (2024). Will artificial intelligence replace the need for second brains entirely?

[Gentner(1983)] Gentner, D. (1983). Structure-mapping: A theoretical framework for analogy. *Cognitive Science*, 7(2), 155–170.

[Gerver et al.(2023)] Gerver, H., et al. (2023). Memory and creativity: A meta-analytic examination of the relationship. *Psychology of Aesthetics, Creativity, and the Arts.*

[Hsieh et al.(2024)] Hsieh, C.-P., et al. (2024). Ruler: What's the real context size of your long-context language models? *arXiv preprint arXiv:2404.06654.*

[Karpathy(2026)] Karpathy, A. (2026). LLM Wiki. *GitHub Gist*, April 4, 2026.

[Kiewra(1989)] Kiewra, K. A. (1989). A review of note-taking: The encoding-storage paradigm and beyond. *Educational Psychology Review*, 1(2), 147–172.

[Li et al.(2024)] Li, Z., et al. (2024). Retrieval augmented generation or long-context LLMs? A comprehensive study and hybrid approach. *Proceedings of EMNLP 2024 (Industry Track).*

[Lewis et al.(2024)] Lewis, M., Mitchell, T., et al. (2024). Testing LLMs on structural analogy tasks. *Proceedings of the Annual Meeting of the Cognitive Science Society.*

[Liu et al.(2024)] Liu, N. F., et al. (2024). Lost in the middle: How language models use long contexts. *Transactions of the Association for Computational Linguistics*, 12, 157–173.

[Mark et al.(2005)] Mark, G., Gonzalez, V. M., & Harris, J. (2005). No task left behind? Examining the nature of fragmented work. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 321–330.

[McKinsey(2012)] McKinsey Global Institute. (2012). *The social economy: Unlocking value and productivity through social technologies.*

[Mednick(1962)] Mednick, S. A. (1962). The associative basis of the creative process. *Psychological Review*, 69(3), 220–232.

[Microsoft(2025)] Microsoft. (2025). *Microsoft Work Trend Index 2025.*

[MITRE(2009)] MITRE. (2009). Factors impeding wiki use in the enterprise: A case study.

[Nature(2026)] Nature. (2026, April 1). Hallucinated citations are polluting the scientific literature. *Nature.*

[Nesbit & Adesope(2006)] Nesbit, J. C., & Adesope, O. O. (2006). Learning with concept and knowledge maps: A meta-analysis. *Review of Educational Research*, 76(3), 413–448.

[NIST(2024)] NIST. (2024). *Artificial Intelligence Risk Management Framework: Generative Artificial Intelligence Profile (NIST AI 600-1).*

[Noy & Zhang(2023)] Noy, S., & Zhang, W. (2023). Experimental evidence on the productivity effects of generative artificial intelligence. *Science*, 381(6654), 187–192.

[Peters & Chin-Yee(2025)] Peters, U., & Chin-Yee, B. (2025). LLM summarization overgeneralization. *Royal Society Open Science.*

[Risko & Gilbert(2016)] Risko, E. F., & Gilbert, S. J. (2016). Cognitive offloading. *Trends in Cognitive Sciences*, 20(9), 676–688.

[Yoghourdjian et al.(2019)] Yoghourdjian, V., et al. (2019). Exploring the limits of complexity: A survey of empirical studies on graph visualisation. *Visual Informatics*, 3(4), 264–282.