

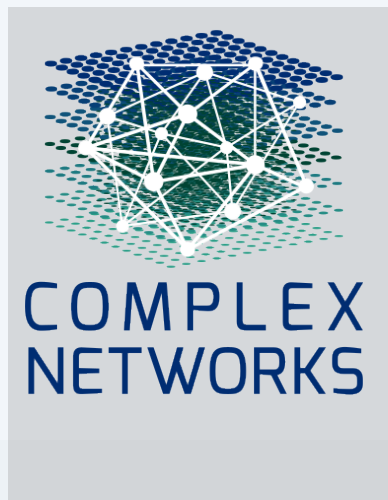
Cognitive MRI of AI Conversations: A Single-User Case Study

Revealing the Hidden Topology of Thought

Alex Towell John Matta

Southern Illinois University Edwardsville

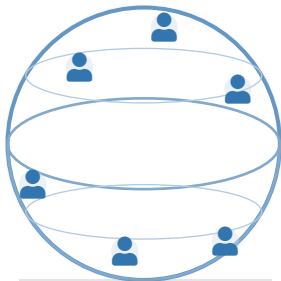
`{atowell,jmatta}@siue.edu`



 Code & Data: github.com/queelius/chatgpt-complex-net

Why Now? The Scale of the Opportunity

1.7 Billion ChatGPT Users



Global Scale

Unprecedented access to
cognitive processes

Why This Matters

Dataset Type		Scale	Process?
Citation Networks		10^8 papers	No
Social Networks		10^9 users	No
LLM	Conversa-	10^9 users	✓
tions			

First-Time Opportunity

Traditional datasets capture **outputs** (papers, posts).
LLM logs capture the **iterative reasoning process** at global scale.

The Big Picture: Externalized Cognition

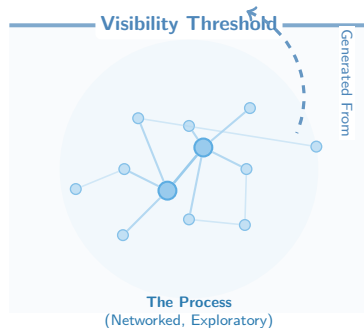
AI conversations are not just chat logs.

We view them through the lens of **Distributed Cognition**:

- **Thinking Out Loud:** The user offloads cognitive load to the machine.
- **The Iterative Loop:** Ideas aren't just “retrieved”; they are constructed through dialogue.

The “Cognitive Dark Matter”

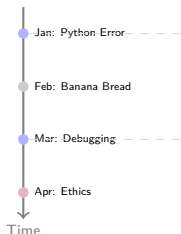
Standard archives preserve the *result* (the paper). LLM logs capture the *process*—the false starts, synthesis, reasoning.



From Log to a “Cognitive MRI”

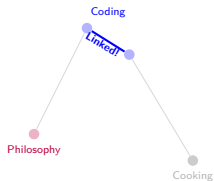
Transforming Linear Time into Semantic Topology

1. The Linear Log (Chronological Sequence)



→
Embed
& Link

2. The Cognitive MRI (Semantic Topology)



The Insight

Distance in **Time** \neq Distance in **Thought**.

The network reconnects ideas (e.g., two coding sessions months apart) that the linear log separates.

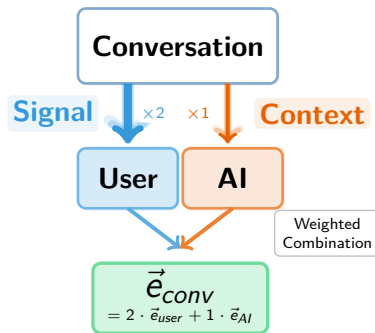
Method: Capturing Intent

The Challenge: AI responses are verbose and generic.

The Solution: Focus on the human.

The “Signal” is the User

- We separate **User Prompts** from **AI Replies**.
- **Weighting:** We prioritize the user's voice (2:1 ratio).
- **Result:** The network connects conversations based on *your* intent, not the AI's boilerplate.

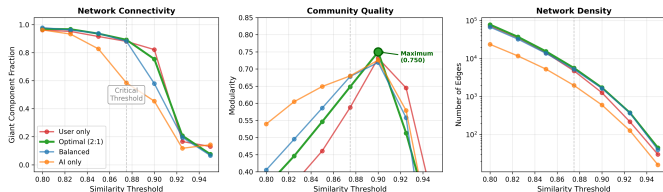


*Embeddings generated via nomic-embed-text (8k context).

Rigorous Parameter Tuning: 2D Ablation Study

We ran a 63-configuration parameter sweep to maximize *Modularity (Q)*.

Phase Transition at $\theta \approx 0.875$: Universal Across All Weight Ratios



Two-Dimensional Sweep

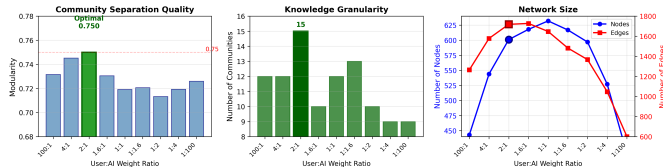
1 Threshold (θ):

- ▶ Phase transition at $\theta = 0.875$
- ▶ **Choice:** $\theta = 0.9$ (optimizes modularity)
- ▶ Below: hairball; Above: fragmentation

2 Weight Ratio (α):

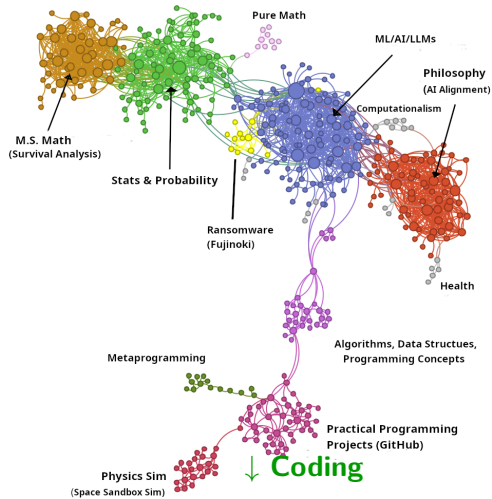
- ▶ Peak at $\alpha = 2 : 1$ (user voice prioritized)
- ▶ **Result:** $Q = 0.750$

Weight Ratio Effects at Threshold $\theta = 0.9$



Data-driven validation of design choices.

The Cognitive MRI: 15 Knowledge Domains



AI Theory →

Giant Component (single user):

- Nodes: 449
- Edges: 1,615
- Modularity: 0.750
- Communities: 15

Clusters emerged
organically – no manual
categorization.

Insight 1: Structural Heterogeneity

Knowledge isn't uniform. Theoretical and practical thinking have distinct shapes.

Theoretical Domains

(Math, Philosophy, ML Theory)



“Small-World” Structures

- **Dense Clustering** ($C \approx 0.58$): Concepts are highly interconnected.
- **Recursive**: Frequent backtracking to refine core definitions (e.g., axioms, ethics).

Practical Domains

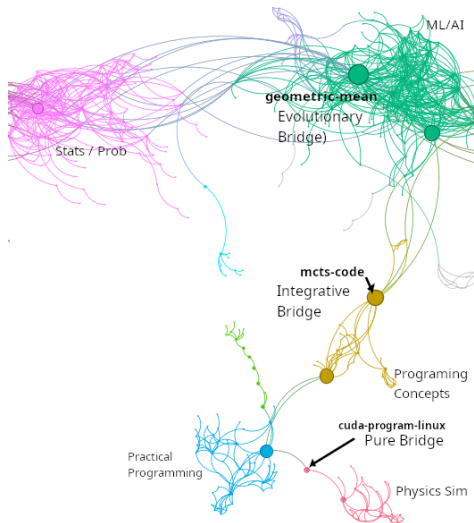
(Programming Projects, Debugging)



Tree-Like Expansion

- **Branching** ($C \approx 0.39$): Task-based exploration without backtracking.
- **Independent**: Projects form isolated silos (e.g., *Metaprogramming* vs. *Physics Sim*).

Insight 2: A Taxonomy of Bridges



The network reveals three distinct bridging mechanisms.

1. Evolutionary Bridges

(e.g., Geometric Mean)

Conversations that **drift** from one topic to another
(Math \rightarrow AI).

2. Integrative Bridges

(e.g., AI Ethics)

Deliberate synthesis of two fields.

3. Pure Bridges

(e.g., CUDA Linux)

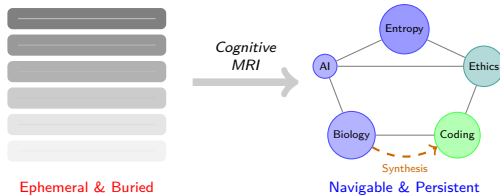
Cognitive Wormholes. Rare shortcuts through conceptual space (e.g., Gaming \leftrightarrow Coding).

The Vision: Personal Knowledge Cartography

Why do we need this map?

Current State
"The Scroll"

Future State
"The Map"



Example Query: "Show me everywhere I discussed entropy."

Result: Network lights up connections: Biology ↔ AI Theory ↔ Coding ↔ Ethics

Problem: Insights buried
in infinite scroll

Solution: Navigate & synthesize
across your entire history

Cognitive MRI: A Proof of Concept

Key Findings



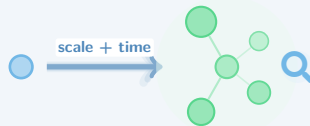
- **Method:** User-weighted embedding + Adaptive Thresholding.
- **Topology:** Heterogeneous (Hubs vs. Trees).
- **Bridges:** Evolutionary, Integrative, & Pure.

Limitations



- Single User & Platform.
- Snapshot in time.
- Exploratory (No “Ground Truth”).

Future Directions



- **Scale:** Multi-user cohorts & cross-platform analysis.
- **Longitudinal:** Track knowledge evolution over time.
- **Validation:** Permutation tests, retrieval benchmarks, user studies.

Backup: Technical Details

Embedding Details


- Model: nomic-embed-text (8k context window)
- Dimension: 768
- Chunking: 500-token windows with 50-token overlap
- User-to-AI weighting: 2:1 ratio (validated via ablation study)

Community Detection

- Algorithm: Louvain (resolution = 1.0)
- Modularity: $Q = 0.750$ (15 communities discovered)
- Giant component: 449 nodes, 1,615 edges

Dataset Filtering

- Original dataset: 1,908 conversations (2 years)
- After similarity threshold ($\theta = 0.9$): 449 conversations in giant component
- Isolated nodes filtered: conversations with no semantic neighbors

Full methodology & code:  github.com/queelius/chatgpt-complex-net

Backup: Core Formulas

Weighted Embedding

$$\vec{e}_{conv} = \frac{\alpha \vec{e}_{user} + \vec{e}_{AI}}{\|\alpha \vec{e}_{user} + \vec{e}_{AI}\|}$$

$\alpha = 2$ (2:1 weighting)

Modularity (Newman's Q)

$$Q = \frac{1}{2m} \sum_{ij} \left[A_{ij} - \frac{k_i k_j}{2m} \right] \delta(c_i, c_j)$$

A_{ij} : adjacency, k_i : degree, m : edges

Betweenness Centrality

$$B(v) = \sum_{s \neq v \neq t} \frac{\sigma_{st}(v)}{\sigma_{st}}$$

σ_{st} : shortest paths $s \rightarrow t$

Clustering Coefficient

$$C_i = \frac{2e_i}{k_i(k_i - 1)}$$

e_i : edges among neighbors

Backup: Privacy & Data Handling

This Study

- **Consent:** Author's own conversations
- **Export:** Official ChatGPT data export
- **Content:** Exploratory/academic only
- **Sharing:** Aggregated statistics, no raw logs

Code Release

- Framework is open-source
- Users run locally on their own data
- No data leaves user's machine

Future Multi-User Studies

- **IRB Required:** Formal ethics review
- **Informed Consent:** Explicit opt-in
- **Anonymization:**
 - ▶ Remove PII (names, emails)
 - ▶ Hash conversation IDs
 - ▶ Redact sensitive topics
- **Differential Privacy:** For aggregate statistics

Key Principle

Designed for **self-knowledge**—users mapping their own thought, not surveillance.

Backup: Methodology Alternatives

Why These Design Choices?

Choice	Alternative	Why We Chose This
Cosine Similarity	Euclidean Distance	Magnitude-invariant (length \neq relevance)
	Jaccard (set-based)	Semantic continuity, not just keywords
Threshold ($\theta=0.9$)	Soft/fuzzy clustering	Clear community boundaries
	k-NN graph	Ablation validated hard threshold
nomic-embed-text	OpenAI embeddings	Open weights, 8k context, reproducible
	Sentence-BERT	Better long-context handling
2:1 Weighting	Equal (1:1)	AI responses dilute user intent
	User-only	Loses conversational context

All choices validated via 63-configuration ablation study (Slide 6)