

From Episodes to Abstractions: Latent Hierarchical Memory in 1,908 AI Conversations

Alex Towell and John Matta

Southern Illinois University Edwardsville, Edwardsville, IL, USA
queelius@gmail.com, jmatta@siue.edu

ORCID: 0000-0001-6443-9897 (A.T.), 0000-0002-7666-1409 (J.M.)

Abstract. We investigate whether AI conversation archives—externalized records of knowledge exploration through dialogue with large language models—contain hierarchical structure analogous to human semantic memory. From a longitudinal archive of 1,908 ChatGPT conversations spanning 29 months, we extract semantic concepts using LLM-based analysis, embed them in a 768-dimensional semantic space, and apply hierarchical agglomerative clustering to organize them into a four-level hierarchy: 500 *meta-concepts* \rightarrow 50 *themes* \rightarrow 8 *knowledge domains*. The concept co-occurrence network is small-world ($\sigma \approx 6.6$, validated against 100 Erdős-Rényi random graphs), matching the topology reported for human semantic memory networks. Meta-concept vocabulary follows Heaps’ law ($\beta = 0.320$), with growth dynamics distinguishable from both random clustering and a bipartite configuration null preserving degree sequences ($p < 0.001$ for both), indicating genuine categorical structure. The bipartite graph further reveals that 77% of episodes span multiple knowledge domains through shared concepts, and that cross-domain concept sharing is concentrated in specific domain pairs rather than distributed uniformly.

Keywords: semantic memory, hierarchical clustering, knowledge networks, complementary learning systems, AI conversations, small-world networks

1 Introduction

Over the past three years, millions of users have engaged in sustained dialogue with large language models (LLMs), producing conversation archives that constitute an unprecedented record of externalized knowledge exploration [24]. A single user’s ChatGPT history may contain thousands of episodes spanning programming, mathematics, philosophy, and creative writing—each episode a trace of a specific learning or problem-solving encounter. Yet these archives are typically treated as flat, chronological lists, discarding whatever structure might exist in the knowledge they collectively represent.

We ask: *does hierarchical structure emerge from these conversation archives when viewed through the lens of cognitive memory theory?*

Complementary Learning Systems (CLS) theory [13,11] generates a testable prediction: its *hippocampal* system rapidly encodes individual episodes, while a *neocortical* system gradually extracts statistical regularities. Novel experiences should increasingly map to existing categories rather than creating new ones—a phenomenon measurable through vocabulary growth dynamics. We test whether this holds for externalized knowledge archives.

Independently, research on human semantic memory networks has established that conceptual knowledge organizes into *small-world* networks [18,17]: high local clustering with short global path lengths. Steyvers and Tenenbaum [18] showed this topology holds across Roget’s Thesaurus, WordNet, and word association norms [3], with clustering-to-path-length ratios corresponding to σ values of 5.6–15.3 under the later Humphries–Gurney formulation [9].

We apply these cognitive frameworks to a longitudinal archive of 1,908 ChatGPT conversations spanning December 2022 through April 2025. Our approach is bottom-up: we use an LLM to extract fine-grained semantic concepts from each episode, embed them in a shared vector space, and allow hierarchical structure to emerge through geometric clustering. The central contributions are:

1. A bipartite episode–concept graph and four-level hierarchy (1,908 episodes \rightarrow 500 meta-concepts \rightarrow 50 themes \rightarrow 8 domains) with interpretable knowledge categories at each level;
2. Sublinear vocabulary growth (Heaps’ law $\beta = 0.320$), with growth dynamics distinguishable from a random-clustering null model ($p < 0.001$), showing that semantic structure creates meaningful categorical distinctions;
3. Small-world topology ($\sigma \approx 6.6$, against 100 Erdős–Rényi random graphs) in the concept co-occurrence network, consistent with human semantic memory benchmarks;
4. Many-to-many episode–concept associations, where 77% of episodes span multiple knowledge domains—structure invisible to partition-based methods;
5. Size-normalized asymmetric concept sharing between domains, revealing genuine semantic isolation punctuated by specific cross-domain dependencies.

2 Related Work

Complementary Learning Systems. CLS theory [13,11] explains how hippocampal episodic encoding and neocortical consolidation interact. The hippocampal system rapidly encodes individual experiences with pattern separation, while the neocortical system gradually extracts statistical regularities through interleaved replay. A key prediction is that novel experiences increasingly map to existing semantic categories—a form of vocabulary saturation that can be measured through growth dynamics like Heaps’ law [8]. Heaps’ law ($V(n) = K \cdot n^\beta$) was originally formulated for natural language vocabulary growth, where $\beta \approx 0.4$ – 0.6 [12]; lower β indicates stronger consolidation.

Cognitive Network Science. Semantic memory networks exhibit small-world topology [23]—high local clustering with short global path lengths [18,17]. Collins and Loftus [2] proposed spreading activation as the retrieval mechanism; Rosch [16] established hierarchical categorization. De Deyne et al. [3] confirmed small-world properties across 12,000 cues in a word association network. The small-world coefficient $\sigma = (C/C_{\text{rand}})/(L/L_{\text{rand}})$, as formalized by Humphries and Gurney [9], compares observed clustering and path lengths to Erdős-Rényi random graphs with matched node and edge counts. From the clustering and path length ratios reported by Steyvers and Tenenbaum [18], one can compute σ values of 5.6–15.3 across their three semantic networks, supporting their argument that small-world structure is a universal property of semantic memory.

Theme Discovery and Knowledge Extraction. LDA [1], BERTopic [7], and Hierarchical Dirichlet Processes [20] discover latent themes but do not produce named concept nodes linked to episodes in a bipartite graph. LLMs have been used for ontology construction [6], knowledge graphs [15], and Graph RAG [4], but these focus on downstream retrieval rather than analyzing the *structure* of extracted knowledge.

Conversation Corpus Analysis. Large-scale LLM conversation datasets include WildChat [24] (1M ChatGPT interactions), LMSYS-Chat-1M [25] (1M multi-model conversations), and OpenAssistant [10] (161K crowd-sourced messages). These corpora support behavioral analysis and model evaluation but do not examine the semantic structure of individual user archives over time. Our prior work [21] constructed a semantic similarity network from the same archive using Louvain partitioning [5]. The present work differs fundamentally: extracting explicit concepts produces a bipartite graph capturing many-to-many associations that partition-based methods cannot represent.

3 Data and Methods

3.1 Dataset

The dataset comprises 1,908 ChatGPT conversations from a single user’s archive, spanning December 2022 through April 2025 (29 months), containing 35,411 messages (16,503 user, 18,908 assistant). Conversations range from 2 to 416 messages (median 10, mean 18.6), covering statistical methodology, software engineering, AI research, R programming, physics, philosophy, and creative projects across multiple model generations (GPT-3.5 through GPT-4o and o1/o3). The archive represents a naturalistic record of sustained AI-assisted knowledge exploration by a single researcher; it is not crowd-sourced or curated for particular topics.

3.2 Concept Extraction

We extract semantic concepts from each episode using Claude Sonnet 3.5 v2¹ via the Anthropic API. For each episode, the model receives the first 20 messages

¹ Model ID: `claude-3-5-sonnet-20241022`, temperature = 0.

of the conversation (each capped at 500 characters) and a structured prompt requesting 3–10 noun-phrase concepts (2–5 words each) capturing the specific knowledge explored. The prompt instructs: *“Be specific—use ‘Metropolis-Hastings algorithm’ rather than ‘statistics.’ Focus on what knowledge is explored, not how the conversation proceeds.”*

To process all 1,908 conversations efficiently, we deploy 20 parallel extraction agents, each operating independently on a partition of episodes. Because agents run in parallel, they cannot share a running vocabulary; post-hoc deduplication is performed geometrically via hierarchical clustering (§3.4).

3.3 Concept Embedding

Each unique concept is embedded using the `nomic-embed-text` model [14], which produces 768-dimensional vectors (the native output dimension of its BERT-derived architecture; this is not a hyperparameter choice). The model was trained via contrastive learning on diverse text and supports 8,192-token context. Embeddings are L2-normalized to enable cosine similarity comparisons.

3.4 Hierarchical Clustering

We apply hierarchical agglomerative clustering (Ward linkage [22], Euclidean distance) to the concept embedding matrix and cut the dendrogram at three heights, producing four levels:

- **Level 0:** 1,908 episodes (individual conversations)
- **Level 1:** 500 meta-concepts (semantic deduplication)
- **Level 2:** 50 themes (broad topic groups)
- **Level 3:** 8 domains (top-level knowledge areas)

The meta-concept level ($k = 500$) groups semantically equivalent concepts from parallel extraction agents (e.g., “bootstrap confidence intervals” and “BCa bootstrap confidence intervals”). Cut points are chosen for interpretability rather than optimality: silhouette analysis shows monotonically increasing scores for $k \geq 20$ with no distinct peak, indicating the absence of natural cluster boundaries in the embedding space. The sensitivity analysis in Table 3 confirms that qualitative findings (sublinear growth, small-world topology) are robust across $k = 200$ –1,000.

3.5 Bipartite Graph and Co-occurrence Network

We construct a bipartite graph $G = (E \cup C, L)$ where E is the set of 1,908 episodes, C is the set of 500 meta-concepts, and $(e, c) \in L$ whenever episode e was assigned at least one raw concept belonging to meta-concept cluster c . This yields 7,153 bipartite edges (mean 3.75 meta-concepts per episode, compared to 3.96 raw concepts per episode). From this bipartite graph, we derive a weighted co-occurrence network on meta-concepts: two meta-concepts $c_i, c_j \in C$ are linked whenever they co-occur in at least one episode, with edge weight $w(c_i, c_j) = |\{e \in E : (e, c_i) \in L \wedge (e, c_j) \in L\}|$.

4 Results

4.1 Concept Extraction and Deduplication

Extraction yields 7,555 raw concept mentions (3.96 per episode) mapping to 3,517 unique concepts after case-insensitive deduplication. The raw distribution is long-tailed: 79.4% of concepts appear in only one episode. The 500-cluster meta-concept level consolidates this further: singletons drop to <1%, and the mean frequency rises from 2.1 to 14.3 episodes per meta-concept (Table 1, Fig. 1).

Table 1. Effect of meta-concept deduplication on bipartite graph density.

Metric	Raw concepts	Meta-concepts
Unique concepts	3,517	500
Singletons (%)	79.4	< 1
Mean episodes/concept	2.1	14.3
Episode pairs sharing concepts	9,304	165,000

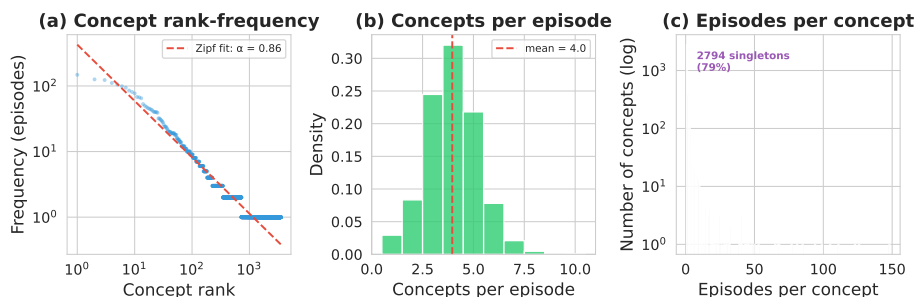


Fig. 1. Concept frequency distributions. Left: Zipf rank-frequency (exponent ≈ 0.86). Center: concepts per episode (mean 3.96). Right: episodes per concept (79% singletons).

4.2 Latent Hierarchy

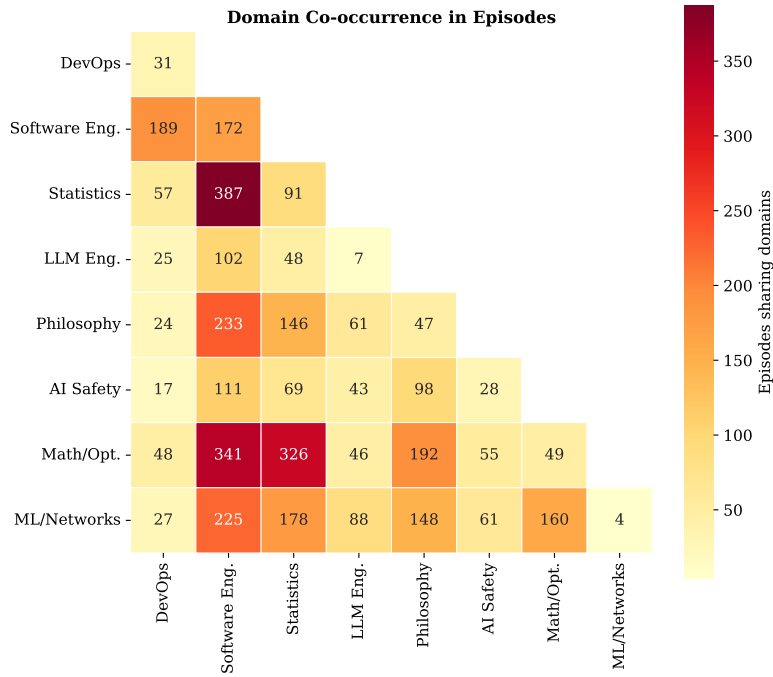
Ward linkage at $k = 8$ produces eight interpretable knowledge domains (Table 2). The bipartite structure allows episodes to span multiple domains: only 22.5% are confined to a single domain; 42% span two, 25% span three, and 10% span four or more. Episodes spanning 3+ domains (35%) function as cross-domain integration points—e.g., “automatic differentiation for MLE in R” spans Software Engineering, Statistics, and Mathematics (Fig. 2).

4.3 Vocabulary Growth and Null Model

The meta-concept vocabulary grows sublinearly following Heaps’ law $V(n) = K \cdot n^\beta$ with $\beta = 0.320$ (Fig. 3); under 100 random episode orderings, $\beta = 0.241 \pm 0.007$, confirming robustness to ordering. The raw vocabulary grows

Table 2. Eight knowledge domains from Ward linkage clustering.

Domain	Raw concepts	Representative concepts
Software Engineering	1,074	code refactoring, API design
Math & Optimization	634	simulation, Fisher info.
Statistical Methods	510	MLE, bootstrap, Weibull
Philosophy & AI Theory	489	consciousness, complexity
ML & Network Science	274	language models, networks
DevOps & Integration	273	APIs, hosting, admin
AI Safety & Research	177	alignment, evaluation
LLM Engineering	86	prompts, agents, tools

**Fig. 2.** Domain co-occurrence heatmap. Software Engineering co-occurs broadly; specialized pairs (Statistics–Math) show focused affinity.

nearly linearly ($\beta = 0.931$), so sublinear growth arises from clustering, not extraction.

Since fixed $k = 500$ guarantees sublinear growth regardless of content, we test whether the observed β reflects genuine structure using two null models. **Null 1** (cluster permutation): randomly permute concept-to-cluster assignments (preserving sizes, 1,000 permutations); $\beta_{\text{null}} = 0.268 \pm 0.006$, $p < 0.001$. **Null 2** (bipartite configuration): randomize which meta-concepts appear in which episodes while preserving both marginals via the Curveball algorithm [19]; $\beta_{\text{null}} = 0.254 \pm 0.007$, $p < 0.001$. Both confirm that the observed growth dynamics are not ceiling or degree-sequence artifacts. Semantic clustering produces

higher β than either null: coherent clusters create meaningful distinctions, so encountering a new meta-concept requires genuinely novel territory. Sensitivity analysis confirms robustness across $k = 200$ –1,000 (Table 3).

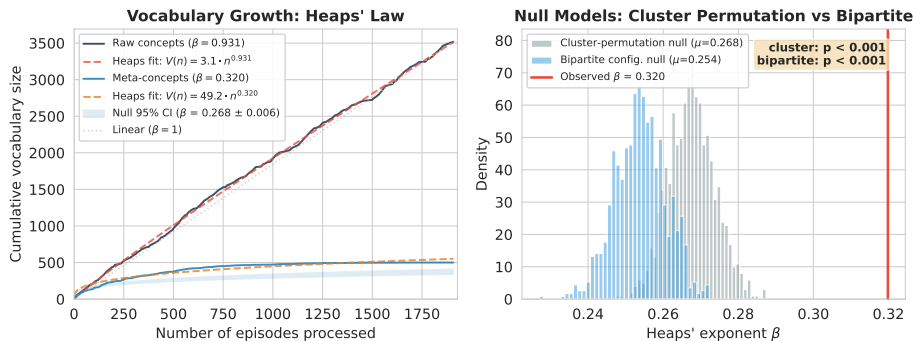


Fig. 3. Left: Heaps' law vocabulary growth ($\beta = 0.931$ raw, 0.320 meta-concept). Right: null model distributions—cluster-permutation ($\beta_{\text{null}} = 0.268$) and bipartite configuration ($\beta_{\text{null}} = 0.254$), both $p < 0.001$.

4.4 Small-World Topology

The meta-concept co-occurrence network (499 of 500 meta-concepts; one singleton has no co-occurrence edges; 5,935 edges) exhibits small-world properties (Table 3). Following the Humphries–Gurney formulation [9], we compare clustering and path length to 100 Erdős–Rényi $G(n, m)$ random graphs with matched node and edge counts. The clustering coefficient is $6.9\times$ higher than random, while the path length is only $1.05\times$ longer, yielding $\sigma \approx 6.6$. This is comparable to σ values derivable from the clustering and path-length ratios reported for human semantic memory networks [18,3]—Roget's ($\sigma \approx 13$), WordNet ($\sigma \approx 15$), and word associations ($\sigma \approx 5.6$)—though direct comparison is approximate since those studies used different random graph baselines. Clustering sensitivity confirms robustness: σ ranges from 2.0 ($k = 200$) to 12.6 ($k = 700$), remaining above 1 across all tested values of k (Table 3).

Table 3. Small-world analysis at $k = 500$ (top) and sensitivity across meta-concept granularity (bottom). Random graph baselines: 100 $G(n, m)$ graphs.

Property	Value	Interpretation	k	Heaps' β	σ	C	Edges
Nodes / Edges	499 / 5,935		50	0.061	1.1	0.830	906
Clustering C	0.329		100	0.103	1.4	0.652	2,286
C/C_{rand}	6.89	High local clustering	200	0.158	2.0	0.425	4,111
L/L_{rand}	1.05	Short global paths	300	0.225	2.9	0.351	5,073
Small-world σ	6.57	Strongly small-world	500	0.320	6.57	0.329	5,935
			700	0.382	12.6	0.351	6,593
			1000	0.467	25.9	0.401	7,291

4.5 Asymmetric Concept Flow

The bipartite graph enables a directed analysis that the undirected co-occurrence network cannot capture: for each pair of domains (A, B) , we measure the fraction of B 's episodes that contain at least one concept from domain A . We call this fraction the *broadcast reach* of domain A —the share of non- A episodes containing at least one of A 's concepts. Conversely, the *porosity* of domain A is the fraction of A 's episodes containing at least one foreign concept. If concept sharing were symmetric, broadcast reach from A to B would equal the reverse; in practice, it is asymmetric (Fig. 4).

Software Engineering has the highest broadcast reach (40% of all non-SE episodes), penetrating all seven other domains. However, Software Engineering also contains 30.5% of all raw concepts (1,074 of 3,517), raising the question of whether its broad reach reflects genuine cross-domain dependency or a domain size effect. To control for this, we compare observed flow to a null model that randomly reassigns concept-domain labels (preserving domain sizes, 100 permutations). The observed/expected ratios fall below 1.0 for most domain pairs (mean = 0.61), confirming that the eight domains capture genuine semantic boundaries: concepts cluster within domains more tightly than expected under random concept-domain assignment.

Within this overall isolation, specific domain pairs show elevated sharing relative to the null. LLM Engineering \rightarrow ML & Network Science has an observed/expected ratio of 2.67, and LLM Engineering \rightarrow AI Safety reaches 1.27. ML & Network Science has the highest porosity (30% of its episodes contain foreign concepts), functioning as an interdisciplinary junction where concepts from multiple domains converge.

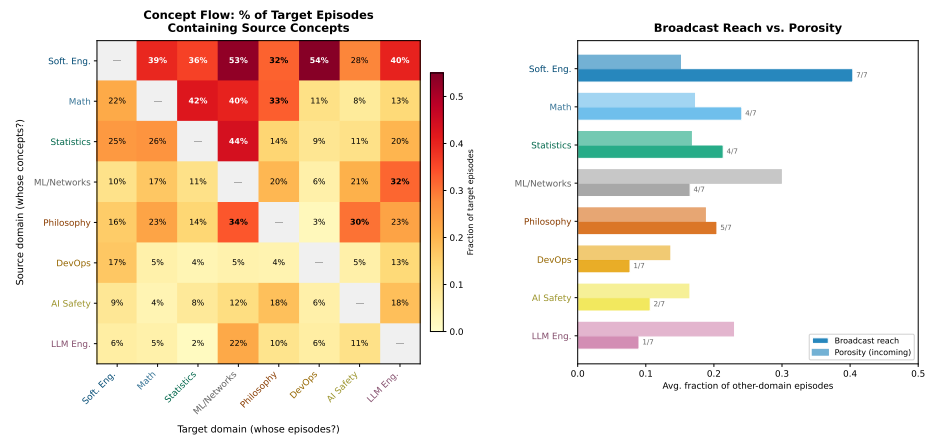


Fig. 4. Asymmetric concept flow. Left: fraction of target-domain episodes (columns) containing source-domain concepts (rows). Right: broadcast reach vs. porosity per domain; numbers indicate reach breadth ($>15\%$).

4.6 Bipartite Network Visualization

Figure 5 visualizes the full bipartite structure as a radial layout (meta-concepts inner ring, episodes outer ring, grouped by domain). Cross-sector edges make the asymmetric flow visible: Software Engineering sends concept links into every sector, while Philosophy connects almost exclusively within its own domain.

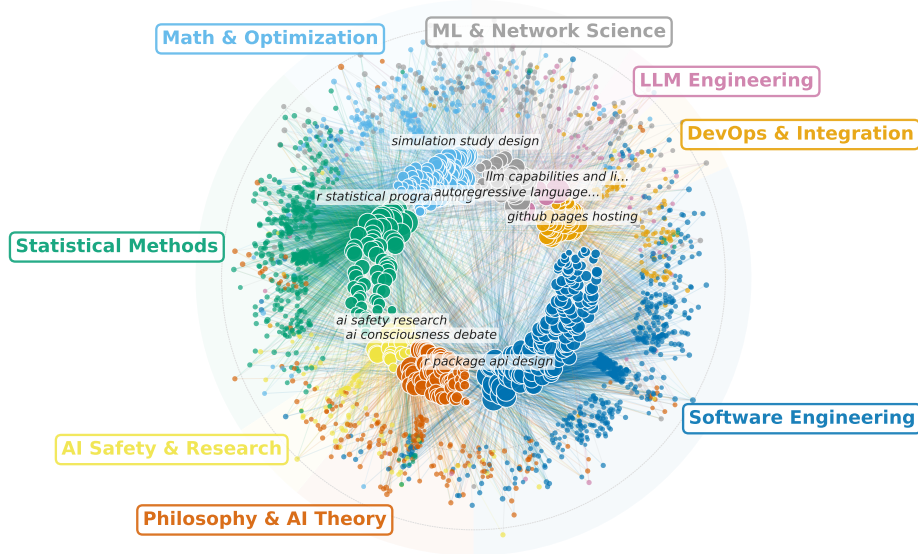


Fig. 5. Bipartite knowledge map. Inner ring: 499 meta-concepts; outer ring: 1,908 episodes (both colored by domain). Cross-sector edges reveal directed concept flow: foundational domains export broadly, applied domains absorb.

Figure 6 illustrates how two episodes from different domains—“AI Language Models Compared” (Statistical Methods) and “Language Models and Computation” (Philosophy & AI Theory)—share five meta-concepts spanning four domains through shared concepts like Solomonoff induction and Kolmogorov complexity. A partition-based method [5] would assign them to separate communities; the bipartite hierarchy captures their relationship explicitly.

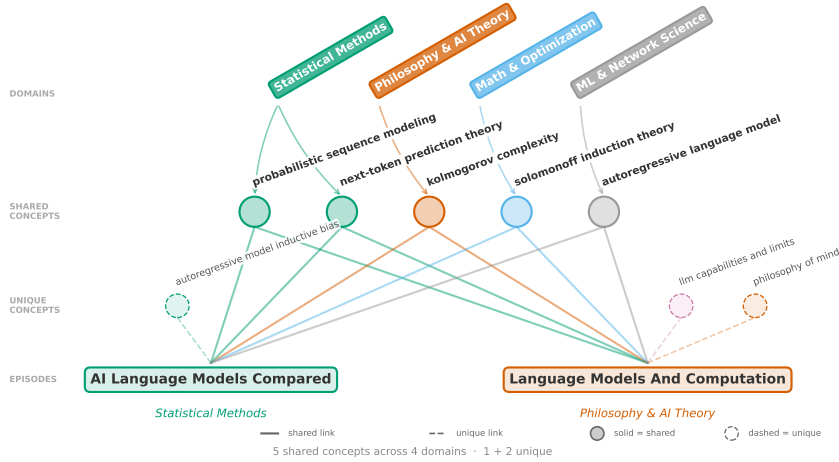


Fig. 6. Two episodes from different domains sharing 5 meta-concepts across 4 domains. Solid: shared concepts; dashed: unique to each episode.

5 Discussion

Vocabulary Growth and Semantic Structure. Since clustering at fixed k guarantees sublinear growth, the absolute β is not evidence of consolidation. The null model comparisons are therefore essential: semantic clustering produces *higher* β than both random cluster assignment ($\beta_{\text{null}} = 0.268$) and the bipartite configuration null preserving degree sequences ($\beta_{\text{null}} = 0.254$), both $p < 0.001$. Under either null, unrelated items share clusters by chance; under semantic clustering, only genuinely related episodes share a meta-concept, so encountering a new category requires exploring novel territory. CLS theory predicts that meaningful categories *facilitate* consolidation (lower β); we observe the opposite. An alternative reading—that the clustering over-distinguishes—is less likely given that the pattern holds across $k = 200$ – $1,000$ (Table 3). The archive’s categorical structure maintains discriminable representations, consistent with preserving distinctions rather than collapsing them.

Small-World Topology and Comparison with Human Memory Networks. The small-world coefficient $\sigma \approx 6.6$ falls within the range reported for human semantic memory networks ($\sigma \approx 5$ – 15) [18]. We do not claim conversation archives *are* semantic memory, but this structural convergence suggests they function as an externalized analog exhibiting the same balance of local coherence and global accessibility. The *qualitative* conclusion (small-worldness, $\sigma > 1$) is robust across $k = 100$ – $1,000$ (Table 3), though the *quantitative* comparison to human benchmarks depends on k .

Many-to-Many Associations and Cross-Domain Flow. The bipartite graph captures knowledge polysemy invisible to partition-based methods [5]: 77% of episodes span multiple domains. Size-normalized flow confirms genuine semantic boundaries (mean observed/expected = 0.61) punctuated by targeted dependencies

(LLM Engineering \rightarrow ML & Network Science at $2.67\times$ expected). The bipartite structure and domain hierarchy could enable concept-level indexing for Graph RAG [4].

Limitations. This single-user proof-of-concept requires replication on multi-user corpora [24,25]. Hierarchy cut points are chosen for interpretability (no natural cluster boundaries), and concept extraction depends on the LLM used. Despite this, convergence across independent metrics suggests general features of categorical knowledge systems.

6 Conclusion

We have shown that 1,908 AI conversations contain latent hierarchical structure: a four-level hierarchy with interpretable domains, sublinear vocabulary growth validated against dual null models (both $p < 0.001$), and small-world topology ($\sigma \approx 6.6$) matching human semantic memory benchmarks. The bipartite structure captures cross-domain integration invisible to partition-based methods, while size-normalized flow reveals genuine semantic boundaries. These results position AI conversation archives as structured knowledge artifacts whose organization exhibits structural parallels with human semantic memory networks.

Data and Code Availability.

Code: <https://github.com/queelius/chatgpt-complex-net> (DOI: 10.5281/zenodo.15314235).

References

1. Blei, D.M., Ng, A.Y., Jordan, M.I.: Latent Dirichlet allocation. *Journal of Machine Learning Research* **3**, 993–1022 (2003)
2. Collins, A.M., Loftus, E.F.: A spreading-activation theory of semantic processing. *Psychological Review* **82**(6), 407–428 (1975)
3. De Deyne, S., Navarro, D.J., Perfors, A., Brysbaert, M., Storms, G.: The “Small World of Words” English word association norms for over 12,000 cue words. *Behavior Research Methods* **51**(3), 987–1006 (2019). DOI 10.3758/s13428-018-1115-7
4. Edge, D., Trinh, H., Cheng, N., Bradley, J., Chao, A., Mody, A., Truitt, S., Larson, J.: From local to global: A Graph RAG approach to query-focused summarization (2024)
5. Fortunato, S.: Community detection in graphs. *Physics Reports* **486**(3–5), 75–174 (2010). DOI 10.1016/j.physrep.2009.11.002
6. Funk, M., Hosemann, S., Jung, J.C., Lutz, C.: Towards ontology construction with language models. In: Knowledge Base Construction from Pre-Trained Language Models Workshop (KBC-LM) at ISWC (2023)
7. Grootendorst, M.: BERTopic: Neural topic modeling with a class-based TF-IDF procedure. *arXiv preprint arXiv:2203.05794* (2022)
8. Heaps, H.S.: *Information Retrieval: Computational and Theoretical Aspects*. Academic Press (1978)

9. Humphries, M.D., Gurney, K.: Network ‘small-world-ness’: A quantitative method for determining canonical network equivalence. *PLoS ONE* **3**(4), e0002,051 (2008). DOI 10.1371/journal.pone.0002051
10. Köpf, A., Kilcher, Y., von Rütte, D., Anagnostidis, S., Tam, Z.R., Stevens, K., Barhoum, A., Duc, N.M., Stanley, O., Nagyfi, R., et al.: OpenAssistant conversations—democratizing large language model alignment. In: *Advances in Neural Information Processing Systems*, vol. 36 (2023)
11. Kumaran, D., Hassabis, D., McClelland, J.L.: What learning systems do intelligent agents need? Complementary learning systems theory updated. *Trends in Cognitive Sciences* **20**(7), 512–534 (2016)
12. Manning, C.D., Raghavan, P., Schütze, H.: Heaps’ law: Estimating the number of terms. In: *Introduction to Information Retrieval*, chap. 5, pp. 86–87. Cambridge University Press, Cambridge, UK (2008)
13. McClelland, J.L., McNaughton, B.L., O’Reilly, R.C.: Why there are complementary learning systems in the hippocampus and neocortex: insights from the successes and failures of connectionist models of learning and memory. *Psychological Review* **102**(3), 419–457 (1995)
14. Nussbaum, Z., Morris, J.X., Duderstadt, B., Mulyar, A.: Nomic embed: Training a reproducible long context text embedder. arXiv preprint arXiv:2402.01613 (2024)
15. Pan, S., Luo, L., Wang, Y., Chen, C., Wang, J., Wu, X.: Unifying large language models and knowledge graphs: A roadmap. *IEEE Transactions on Knowledge and Data Engineering* **36**(7), 3580–3599 (2024)
16. Rosch, E.: Cognitive representations of semantic categories. *Journal of Experimental Psychology: General* **104**(3), 192–233 (1975)
17. Siew, C.S.Q., Wulff, D.U., Beckage, N.M., Kenett, Y.N.: Cognitive network science: A review of research on cognition through the lens of network representations, processes, and dynamics. *Complexity* **2019**, 2108,423 (2019). DOI 10.1155/2019/2108423
18. Steyvers, M., Tenenbaum, J.B.: The large-scale structure of semantic networks: statistical analyses and a model of semantic growth. *Cognitive Science* **29**(1), 41–78 (2005)
19. Strona, G., Nappo, D., Blber, F., Fattorini, S., San-Miguel-Ayanz, J.: A fast and unbiased procedure to randomize ecological binary matrices with fixed row and column totals. *Nature Communications* **5**, 4114 (2014). DOI 10.1038/ncomms5114
20. Teh, Y.W., Jordan, M.I., Beal, M.J., Blei, D.M.: Hierarchical Dirichlet processes. *Journal of the American Statistical Association* **101**(476), 1566–1581 (2006)
21. Towell, A., Matta, J.: Cognitive MRI of AI conversations: Semantic network analysis of a ChatGPT conversation archive. In: *Complex Networks & Their Applications XIII*, pp. 328–339. Springer (2025)
22. Ward, J.H.: Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association* **58**(301), 236–244 (1963)
23. Watts, D.J., Strogatz, S.H.: Collective dynamics of ‘small-world’ networks. *Nature* **393**(6684), 440–442 (1998)
24. Zhao, W., Ren, X., Hessel, J., Cardie, C., Choi, Y., Deng, Y.: WildChat: 1M ChatGPT interaction logs in the wild (2024)
25. Zheng, L., Chiang, W.L., Sheng, Y., Li, T., Zhuang, S., Wu, Z., Zhuang, Y., Li, Z., Lin, Z., Xing, E.P., Gonzalez, J.E., Stoica, I., Zhang, H.: LMSYS-Chat-1M: A large-scale real-world LLM conversation dataset. In: *International Conference on Learning Representations* (2024)