Slide 1: Cognitive MRI of AI Conversations

"Good morning. This talk is about a simple question: what if you treat your AI chat history as data?"

"Most of us have hundreds, even thousands, of conversations with ChatGPT and other AI assistants by now. And they just sit there, buried in a timeline. We wanted to know: is there structure hiding in there?"

"We call this a 'Cognitive MRI.' That's a metaphor—we're trying to extract a kind of knowledge map from chat logs."

Slide 2: Why Now? The Scale of the Opportunity

"ChatGPT has 1.7 billion users. That's a lot of conversations."

"And here's what's interesting. These logs capture something different. Citation networks show you papers—outputs. Social networks show you who knows whom—connections."

"But chat logs might show you the *process*. How ideas develop. The back-and-forth."

"So our question: can we extract a meaningful knowledge map from someone's chat history? Today I'll show you one case study."

Slide 3: The Big Picture: Externalized Cognition

"There's a concept called Distributed Cognition. The idea is simple: thinking doesn't just happen in your head. It happens between you and your tools."

"When you talk to an LLM, you're thinking out loud. Offloading work. And ideas get built through that conversation—not just looked up."

"Traditionally, what gets archived? The product. The final draft. Clean and polished."

"The messy part—the exploring, the dead ends—that usually disappears."

"Think about fixing a bug. Twenty rounds of trying things. False starts, backtracking, trying something else. That's where you actually learn how the system works. But the commit message? One line: 'fixed bug.' "

"Or a mathematician. Months of redefining the problem. Filling notebooks with attempts that don't work. That struggle is where the insight comes from. But we only ever see the final theorem."

"That's what I call cognitive dark matter. And chat logs might capture it."

"So here's what we did."

Slide 4: From Log to a "Cognitive MRI"

"So how do we turn a linear chat log into a network?"

"Start with a chat history. It's just a timeline. One week you're asking about some Python bug. Next week, a banana bread recipe. Month later, back to debugging. Then something about ethics."

"We turn each conversation into a vector—an embedding. Similar topics land close together in that space. Like word2vec, where 'king' and 'queen' end up near each other. Same idea, but for whole conversations."

"Then we connect by similarity, not by time. If two conversations are similar enough—above some threshold we call theta—we draw an edge."

"What happens? Those two coding chats—months apart—snap together. Banana bread floats off on its own. Ethics clusters somewhere else."

"The point is: close in time doesn't mean close in thought. The network reconnects what the timeline splits apart."

Slide 5: Method: Capturing Intent

"There's a problem though. AI responses are wordy. Lots of filler. If you treat everything the same, the AI's voice drowns out yours."

"So we split user messages from AI messages, and weight them differently. Why? Your prompts carry the intent. What you actually care about. The AI just adds context."

"We take the average of your messages, the average of the AI's, then blend them. One vector per conversation."

"That gives us two knobs to tune. Alpha: how much to weight user versus AI. And theta—the similarity cutoff I mentioned—how close do two conversations need to be before we draw an edge?"

"Both parameters need testing. That's the ablation study."

Slide 6: Rigorous Parameter Tuning: 2D Ablation Study

"We tested 63 combinations—swept through different values of theta and alpha together. We optimized for modularity, which measures how clean the community boundaries are."

"For theta, there's a tipping point around 0.875. Below that, everything connects—it's a mess. Above that, things fall apart. 0.9 hit the sweet spot: clear structure, not too sparse."

"For alpha, 2-to-1 user-to-AI gave the best modularity. That backs up the intuition—your voice matters more than the AI's."

"So our final parameters: theta 0.9, alpha 2-to-1, giving us modularity of 0.750."

"The point is: these choices aren't arbitrary. We tested them. Though it's still just one person's data."

Slide 7: The Cognitive MRI: 15 Knowledge Domains

[Pause—let them look]

"Here's the result. 449 conversations, two years of chats, 1,615 edges connecting them into 15 communities. Modularity of 0.750—that's strong structure, not noise."

"Let me walk you through the map. Up top-left in orange, that's my master's thesis work in pure math. Next to it in green, stats and probability. Then ML, AI, and LLMs in blue. And over here in red, philosophy and AI alignment. You can see how they flow into each other."

"Down here, coding projects—more spread out. Different projects form their own clusters. There are other small clusters scattered around; you can scan those yourself."

"What's interesting is it's not uniform. About a quarter of the network is this dense core that connects everywhere. The outer parts are more specialized. Average path length? About 6 hops between any two conversations."

"I labeled these communities myself—the algorithm finds the patterns, I interpret what they mean. Though I tested having an LLM do it, and it worked reasonably well. You could automate the whole thing."

"The point is: the algorithm found real structure. Whether it generalizes beyond my own head? That's the N-of-1 limitation."

Slide 8: Insight 1: Structural Heterogeneity

"Different topics have different shapes."

"Theory—math, ML concepts, philosophy—clusters tightly. Clustering coefficient around 0.58. That's high. Small-world structure."

"That makes sense. Theory means revisiting core ideas. Refining definitions. Everything links back."

"Coding is looser. Around 0.39. More like a tree."

"You fix one bug, move on. Less backtracking. Projects stay in their own lanes. Metaprogramming here, physics simulation there. Not much crossover."

"This is suggestive. We'd need more users to know if it holds up."

Slide 9: Insight 2: A Taxonomy of Bridges

"We also looked at the connectors—the high-betweenness nodes that link different clusters."

"Three patterns showed up."

"Evolutionary bridges. Conversations that drift. See 'Geometric Mean' here—the title sounds narrow, right? But look at the actual conversation and it's sprawling. Started in pure math, wandered through probability, ended up in neural networks. The title only captures where it began."

"Integrative bridges. Deliberate synthesis. 'mcts-code' here—Monte Carlo Tree Search implementation. It explicitly combines ML and AI concepts with practical programming. You're consciously building connections between theory and practice."

"Pure bridges. Rare but powerful. 'cuda-program-linux'—a single Linux config question that happens to link my physics space sandbox with practical programming. A shortcut through conceptual space."

"We're proposing this as a taxonomy. Based on what we saw."

Slide 10: The Vision: Personal Knowledge Cartography

"So why does any of this matter?"

"Right now, your chat history is organized by time. There's keyword search, sure. But you need to remember what words you used. You're searching by date, hoping you remember a phrase."

"But what if you could search by topic instead? Query: 'Show me everything about entropy.' And the network lights up—biology connects to AI connects to coding connects to ethics. You see the whole landscape at once."

"Now, this paper is about the structure itself—communities, bridges, topology. But once you have that structure, it enables other things. Semantic search. Recommendations. Finding gaps in your knowledge. We haven't built those yet."

"The problem is insights buried in a timeline. The solution is a map. That's where this is heading."

Slide 11: Cognitive MRI: A Proof of Concept

"So what did we show? That LLM conversation logs can be transformed into meaningful cognitive maps. The structure is real—not noise."

"The key contributions: user-weighted embeddings to capture intent. We found heterogeneous topology—theoretical domains form hubs, practical domains form trees. We proposed a taxonomy of three bridge types. And we validated all of this through a rigorous 2D ablation study of 63 configurations."

"But let's be honest about the limitations. This is N-of-1—one user, one platform, one snapshot in time. No ground truth to validate against. We don't know yet if these patterns generalize."

"Where do we go from here? More users. Track how these networks evolve over time. User studies to see if the map actually helps people find things."

"If you have access to larger datasets, we'd love to collaborate."

"Thanks. Happy to take questions."