

# From Storage to Interpretation

## User Perceptions, Practices, and Challenges with Agent Memory

Brennan Jones<sup>1,2</sup>, Nazar Ponochevnyi<sup>1</sup>, Kelsey Stemmler<sup>1</sup>, Emily Su<sup>1</sup>, Young-Ho Kim<sup>3</sup>, Anastasia Kuzminykh<sup>1</sup>

brennan.jones@xjtu.edu.cn, nazar@cs.toronto.edu, kelsey@stemmler.tech, em.su@mail.utoronto.ca, yghokim@younghokim.net, anastasia.kuzminykh@utoronto.ca

<sup>1</sup>Faculty of Information (iSchool), University of Toronto

<sup>2</sup>Department of Computing, Xi'an Jiaotong-Liverpool University

<sup>3</sup>NAVER AI Lab



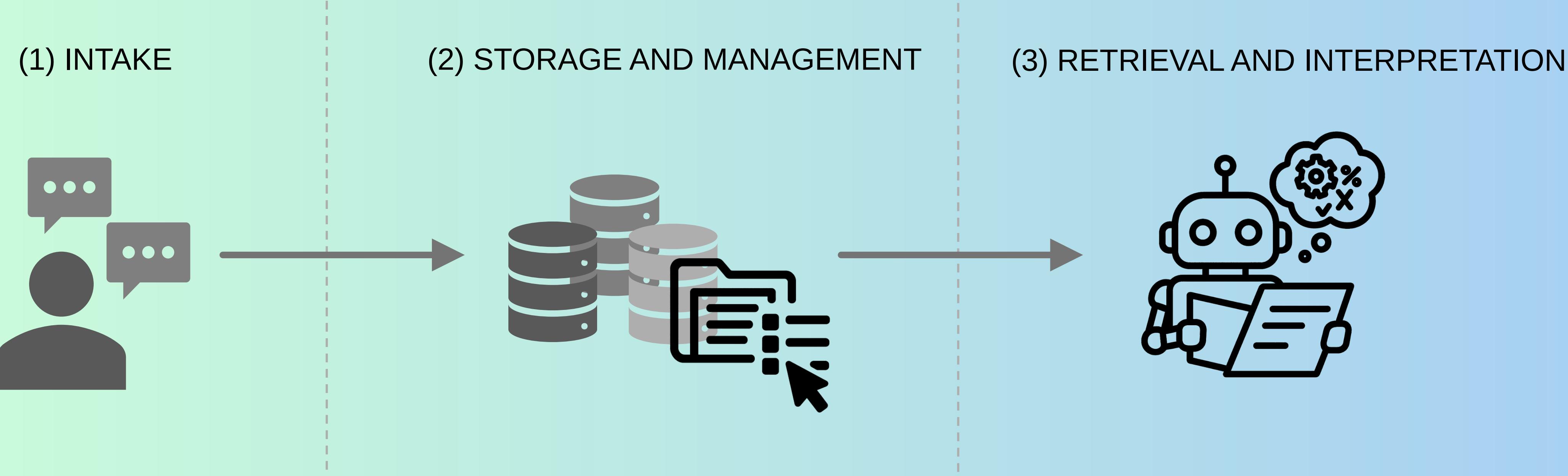
Paper



### Motivation

- People are increasingly interacting with AI agents (e.g., ChatGPT) over long periods of time.
- Such agents can provide always-available personalized and contextualized assistance to users.
- However, in order to achieve this, the user and the agent can benefit from achieving co-understanding of each other, and this relies on the user having proper understanding and influence of the agent's *long-term memory* (LTM).
- While memory in agents has been extensively covered by prior research, there is a lack of understanding of users' perceptions, practices, and challenges with agent LTM. Understanding these could help designers build agent LTM mechanisms that involve users and consider their mental models.

### FINDINGS



Users described agent memory as a process with three interdependent stages.

**Stage 1:** Discrete data (e.g., messages) come in as input from the user. **Stage 2:** Some of this discrete data is stored into the agent's LTM. **Stage 3:** In the future, this information is retrieved from LTM and interpreted by the agent in a new context.

Participants' perceptions were dominated by Stage 3, where meaning is reconstructed. Users judged memory quality by **how well the agent recalled and interpreted information** – not by whether facts were merely retained.

### Research Questions

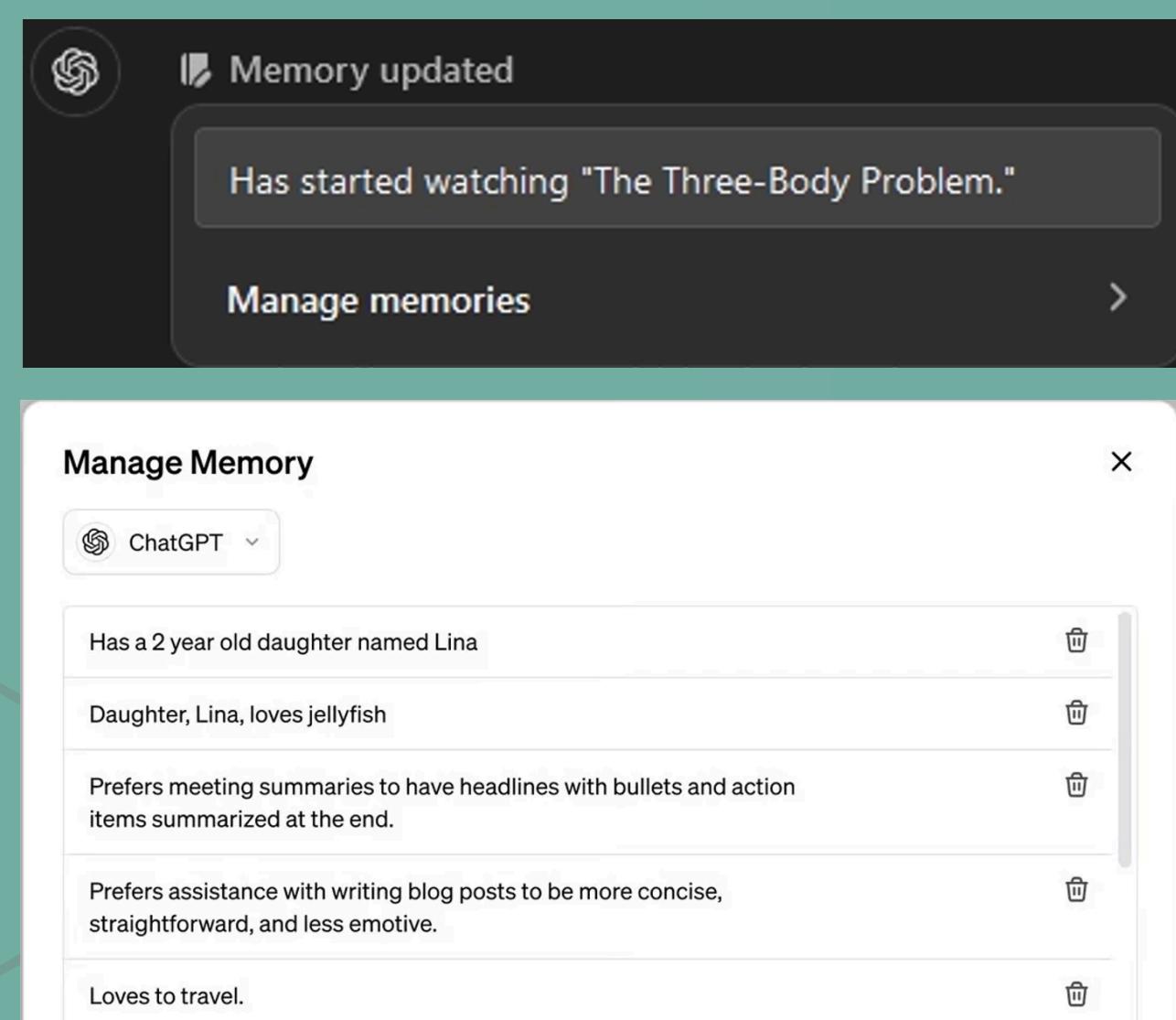
1. What are people's **perceptions and mental models** of agent memory systems?
2. What are people's **current practices** with managing and exploring the memory of an agent?
3. What are people's **challenges and unsatisfied needs** with agent memory?

### Data Collection and Analysis

- We conducted **one-on-one interviews** with **21 participants** who use personalized AI tools with LTM (e.g., ChatGPT, Claude, NotebookLM, Replika) on a regular basis.
- We **thematically coded** the interviews to **better understand users' expectations** of AI systems' LTM mechanisms and their **practices of managing these systems' memories** over longer periods of time.

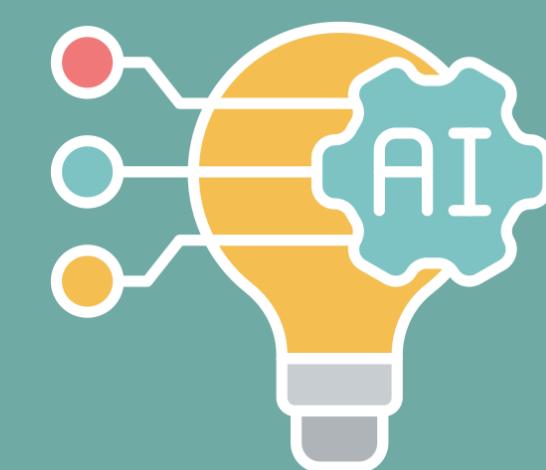
Stage	User's View	Typical Concerns
1 – Intake	Supplying context or prompts	"What does the AI take in?"
2 – Storage & Management	Whether the AI "keeps" info between chats	"Does it save everything?"
3 – Retrieval & Interpretation	How prior info shapes current outputs	"Why did it bring that up?"

### Users' Perceptions



- Many users assume agents either store everything or store nothing in memory.
- Users' perceptions of what the agent does in Stage 3 (*retrieval and interpretation*) are less clear.
  - When the agent fails to provide full transparency, **users develop their own assumptions**.
  - E.g., most participants (P1-P6, P8, P13, P15-P20) assumed that if the interface provided the ability to create multiple chat threads, then the agent was able to **retrieve memories from within the current chat better** than information across chat threads..

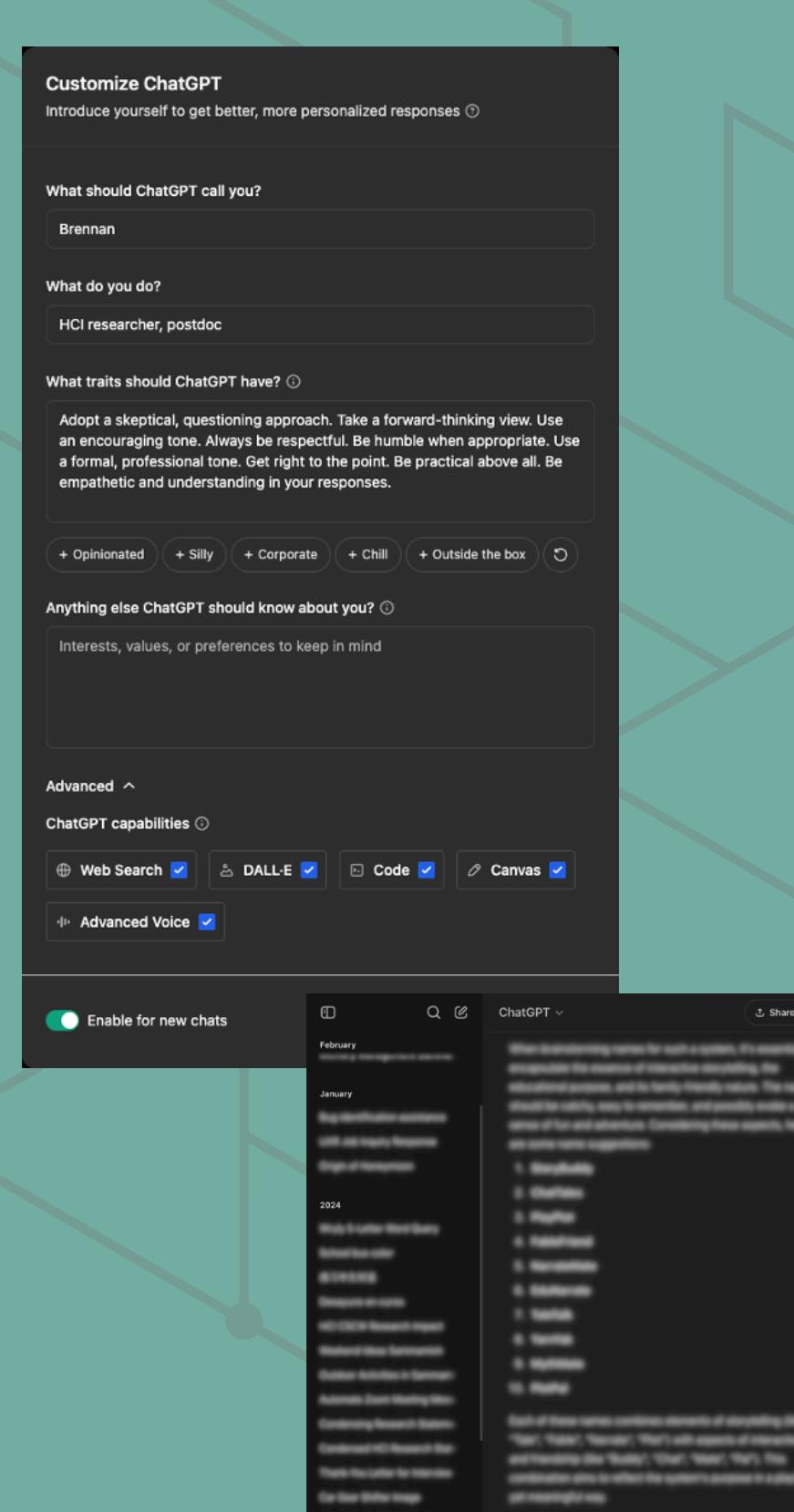
### Users' Practices of Memory Management



Users develop strategies to influence how the agent retrieves and interprets information from LTM, including:

- **Topic separation:** Some participants used multiple chat threads for different projects, tasks, or domains assuming that this keeps memories "separated" or "organized" by topic.
- **Continuity preservation:** Others used a single long thread to maintain flow and shared consistency.
- **Manual transfer:** Some users copied-and-pasted key information or prompts between chat threads, to manually transfer relevant memories from one context to another.
- **Reinforcement cues:** Some users repeated facts or emphasized preferences to 'make sure the agent remembers'.
- **Audit by dialogue:** Some users explicitly queried the agent to see what it remembers and how well (e.g., by asking it "what do you know about X?").

### Users' Challenges and Unsatisfied Needs



Theme	Description	Example User Concern
Opacity of retrieval	Users could see stored memories but not how memories were used.	"It remembers things I didn't mean it to."
Interpretation drift	Agents misapplied past prompts or over-generalized.	Users abandoned sessions when behaviour became odd.
Context leakage	Cross-topic interference between projects or domains.	"Don't mix my research and side projects."
Memory decay	Older context forgotten unless restated.	"I have to pull it back in."
Privacy–utility tension	Users withheld data not for storage fear but for misuse fear.	"I don't mind it knowing; I mind it recalling at the wrong time."

These reflect a consistent desire for retrieval-time transparency and agency: users want to see, edit, and approve how past data is interpreted.

### Key Takeaways

- Users' perceptions of agent LTM are mainly influenced by agents' behaviours, which are affected by **how agents retrieve and interpret information from LTM**.
- Thus, users' practices with managing and exploring LTM are **mainly attempts to influence and understand** how the agent retrieves and interprets memories.
- **Design Recommendation:** interfaces that clarify what the agent remembers, when it retrieves it, and why it applies it.
- **Broader Implication:** Supporting interpretability and selective recall may matter more for user trust than storage capacity itself.