

Responsible NLP Checklist

Paper title: *What Deserves Memory: Adaptive Memory Distillation for LLM Agents*

Authors: *Wenquan Ma, Jiayan Nan, WenLong Wu*

How to read the checklist symbols:

- the authors responded 'yes'
- the authors responded 'no'
- the authors indicated that the question does not apply to their work
- the authors did not respond to the checkbox question

For background on the checklist and guidance provided to the authors, see the [Responsible NLP Checklist](#) page at ACL Rolling Review.

A. Questions mandatory for all submissions.

- A1. Did you describe the limitations of your work?

This paper has a Limitations section.

- A2. Did you discuss any potential risks of your work?

Our work focuses on memory distillation mechanisms for LLM agents and does not introduce novel risks beyond those inherent to LLM-based systems.

B. Did you use or create scientific artifacts? (e.g. code, datasets, models)

- B4. Did you discuss the steps taken to check whether the data that was collected/used contains any information that names or uniquely identifies individual people or offensive content, and the steps taken to protect/anonymize it?

We use publicly available benchmark datasets (LoCoMo, LongMemEval) that have been previously published and vetted by the research community. We did not collect new data.

- B6. Did you report relevant statistics like the number of examples, details of train/test/dev splits, etc. for the data that you used/created?

Section 4.1 (Experimental Setup) reports dataset statistics including number of dialogues/conversations and average token counts for both LoCoMo and LongMemEval benchmarks.

C. Did you run computational experiments?

- C2. Did you discuss the experimental setup, including hyperparameter search and best-found hyperparameter values?

Section 4.1 (Implementation Details) describes all hyperparameters including similarity threshold =0.80, distillation parameters $K_e=K_s=5$, retrieval parameter $k=10$, and model specifications (gpt-4o-mini, gpt-4.1-mini, text-embedding-3-small).

- C3. Did you report descriptive statistics about your results (e.g., error bars around results, summary statistics from sets of experiments), and is it transparent whether you are reporting the max, mean, etc. or just a single run?

We report single-run results following common practice, which is explicitly stated in Section 4.1 (Experimental Setup), Evaluation Metrics, last sentence

The Responsible NLP Checklist used at ACL Rolling Review is adopted from NAACL 2022, with the addition of ACL 2023 question on AI writing assistance and further refinements based on ARR practice. ACL 2026 used a subset of ARR checklist form.

D. Did you use human annotators (e.g., crowdworkers) or research with human subjects?

D1. Did you report the full text of instructions given to participants, including e.g., screenshots, disclaimers of any risks to participants or annotators, etc.?

No human annotators were used in this work

D2. Did you report information about how you recruited (e.g., crowdsourcing platform, students) and paid participants, and discuss if such payment is adequate given the participants' demographic (e.g., country of residence)?

No human annotators were used in this work

D3. Did you discuss whether and how consent was obtained from people whose data you're using/curating (e.g., did your instructions explain how the data would be used)?

No human annotators were used in this work

D4. Was the data collection protocol approved (or determined exempt) by an ethics review board?

No human annotators were used in this work

E. Did you use AI assistants (e.g., ChatGPT, Copilot) in your research, coding, or writing?

E1. If you used AI assistants, did you include information about their use?

AI assistants were used for text editing and code generation, followed by thorough review and verification by the authors. All scientific content, methodology, and experimental design are the original work of the authors. As the AI's role was limited to non-substantive technical assistance, we deemed a formal disclosure within the manuscript unnecessary.