Responsible NLP Checklist

Paper title: *MENTOR: Mitigating Identity Drift in Dynamic Role-Playing via Dual-Chain Structured Memory*
Authors: *Zhu zhuoning, Xingyu Gao, Hailong Shi*

How to read the checklist symbols:

☑ the authors responded 'yes'

☒ the authors responded 'no'

N/A 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?
Ethics Statement and Limitations

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

N/A 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?
Ethics Statement

☑ 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?
6.4 Ablation Study 6 Experiments

☑ **C. Did you run computational experiments?**

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

☑ 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?
6 Experiments and B BEAM-SWITCH Benchmark Details

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

N/A 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. While the exact scoring rubrics and evaluation criteria used by the human annotators are fully detailed in Appendix D (specifically Table 6), we did not include UI screenshots or risk disclaimers.

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.

The annotation task involved evaluating synthetic, role-playing dialogues without exposing annotators to sensitive personal data or potentially harmful/toxic content, thus posing no ethical risks that would necessitate explicit disclaimers.

☒ 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. The human evaluation was a small-scale expert audit (100 samples) conducted internally by graduate students/researchers familiar with the task, rather than a large-scale external crowdsourcing effort. Since external participants were not recruited or financially compensated via crowd-working platforms, detailed recruitment and payment adequacy discussions are not applicable and thus not included.

☐ N/A 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)?
he human evaluation involved auditing synthetic dialogues generated by the system. No sensitive personal data was collected from the annotators, rendering data consent protocols inapplicable.

☐ N/A D4. Was the data collection protocol approved (or determined exempt) by an ethics review board?
Ethics Statement

☒ **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?
Ethics Statement and Limitations