

## Responsible NLP Checklist

Paper title: *IEvoAgent: Evolving Conversational Agent based on User Implicit Feedback*

Authors: *Yichen Cai, Jiayang Li, Junyuan Qiu, Jingya Guo, Weitao You, Changyuan Yang, Lingyun Sun, Pei Chen*

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?

*Yes, the paper discusses technical risks in Section Limitations, and ethical risks in Section Ethical Consideration.*

### 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?

*Yes, the paper discusses these in Section 3.1, Section Ethical Considerations, and Appendix A.6.*

- 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?

*Yes, the paper reports data statistics including turn counts, source distributions, and training/evaluation details across Section 3.1, Section 5.1.*

### C. Did you run computational experiments?

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

*Yes, the paper discusses the experimental setup in Section 5.1 and the hyperparameter values in Section 4 and Appendix A.1.*

- 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?

*Yes, we report average scores in Tables 3-7 and Section 5.3; user study results explicitly specify mean values with statistical testing.*

---

*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.?

*Yes, we provide detail instructions and tasks given to participants in Appendix A.4.*

- 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)?

*Yes, we discuss participant recruitment and compensation in Appendix A.7.*

- 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)?

*Yes, we discuss consent for user study participants in Appendix A.7, and compliance with data usage terms for public datasets in Appendix A.6 and Ethical Considerations.*

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

*Yes, we discuss ethics review board approval for the human subject study in Appendix A.7.*

**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?

*Yes, the paper includes information about the use of AI assistants (specifically Gemini-2.5-Pro) for dataset annotation in Section 3.1 and provides the full text of the specialized prompts used for these AI-based tasks in Appendix A.5.*