

Responsible NLP Checklist

Paper title: *STRIDE-ED: A Strategy-Grounded Stepwise Reasoning Framework for Empathetic Dialogue Systems*

Authors: *Hongru Ji, Yuyin Fan, Meng Zhao, Xianghua Li, Lianwei Wu, Chao Gao*

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?

The Limitations section discusses potential risks of the proposed method, including the lack of statistical validation in data filtering and sampling, as well as possible biases introduced by model selection.

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?

The datasets used in this work are widely used public datasets and contain no personally identifiable information or offensive content, as stated in the Ethical Considerations section.

- 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?

The dataset splits are detailed in Section 4.1, and the dataset statistics are reported in Sections 3.4, 3.5, and 3.6.

C. Did you run computational experiments?

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

Section 4.3 details the experimental setup and reports the optimal values for key hyperparameters.

- 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?

Section 5 states that we have adopted the average value obtained from multiple experiments.

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.?

Sections 4.4 and Ethical Considerations briefly mention our protocols for informing and compensating human evaluators; the full text or screenshots of the instructions are not provided here.

- 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)?

Given the limited scope of the evaluation task, details regarding participant compensation are summarized briefly in Sections 4.4 and the Ethical Considerations section.

- 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)?

The datasets are publicly accessible, with proper citations to the original authors' publications in Section 4.1.

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

The datasets used in this study, detailed in Section 4.1, are well-established and publicly available benchmarks in the empathetic dialogue domain, and their collection protocols comply with standard ethical guidelines, exempt from further institutional review.

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?

The use of AI assistants for our research is comprehensively detailed across the methodology (Sections 3.4 & 3.5), ethical considerations, and the prompt designs in Appendix B.