

Responsible NLP Checklist

Paper title: *Iterative Knowledge Graph Refinement and Integration for Medical Question Answering*

Authors: *Haochen Zou, Yongli Wang*

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?

The potential risks of this work can be found in the Ethical Considerations section.

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 discussion on whether the data contains personally identifiable information or offensive content is provided 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 statistics are reported in the Appendix B Benchmark Datasets section.

C. Did you run computational experiments?

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

The experimental setups are reported in the 4.1 Experimental Setup section, the Appendix D Experimental Configuration section, and the Appendix F Additional Experiments section.

- 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?

The descriptive statistics of the experimental results are reported in the 4.2 Main Results section, the 4.3 Ablation Study section, and the Appendix F Additional Experiments section.

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.?

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.

This work exclusively utilizes existing publicly available benchmark datasets and does not involve human annotators or human participants. No crowdsourcing experiments are conducted in this work that could expose offensive content or collect personally identifiable information. Accordingly, the primary contributions of this work do not involve human subjects.

- 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)?

This work does not include any experiments or detailed information involving participants, nor does it require the recruitment or compensation of human subjects.

- 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)?

This work does not include any experiments, detailed information, or data collection involving crowdsourcing, personal data, or curated information from individuals.

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

This work does not involve human subjects research or any data collection protocols that require approval from an ethics review board.

- 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?

This work does not involve the use of artificial intelligence assistants and strictly adheres to the ACL Code of Ethics, as discussed in the Ethical Considerations section.