

## Responsible NLP Checklist

Paper title: *Knowledge-to-Verification: Exploring RLVR for LLMs in Knowledge-Intensive Domains*

Authors: *Zhonghang Yuan, Zhefan Wang, Fang Hu, Zihong Chen, Jinzhe Li, Gang Li, Jie Ying, Huanjun Kong, Songyang Zhang, Nanqing Dong*

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 paper is mainly centered on presenting the K2V framework and its effectiveness in synthesizing verifiable data to enhance the performance of large language models (LLMs) in knowledge-intensive domains. It does not involve any ethical issues or introduce potential risks.*

### 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 study are publicly available and have been already anonymized by the original publishers, containing no personally identifying information or offensive content that could identify individual people.*

- 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, Appendix A, Appendix B, Appendix C*

### C. Did you run computational experiments?

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

*Section 4.1, Appendix A, Appendix B, Appendix C*

- 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 4, Appendix A*

*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.?

*We recruited human annotators to verify the quality of the knowledge graph (see Appendix G.2). This task did not require the annotators to follow complex instructions. They only needed to make their own judgments. Moreover, this task posed no risks to the annotators.*

- 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)?

*We recruited two PhD students from our lab to carry out the annotation task. They participated voluntarily and chose not to receive any payment. The task is not related to the annotators' demographic, so anyone can participate regardless of their region, gender, or background.*

- 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)?

*(left blank)*

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

*(left blank)*

**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?

*Appendix K*