

## Responsible NLP Checklist

Paper title: *Veri-RI: Toward Precise and Faithful Claim Verification via Online Reinforcement Learning*

Authors: *Qi He, Cheng Qian, Xiusi Chen, Bingxiang He, Yi R. Fung, Heng Ji*

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?

*Training models to enhance their comprehensive verification capability does not exist 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 we used are publicly available benchmark datasets that have been widely adopted in prior work and do not contain personally identifiable information or offensive content.*

- 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?

*We provide details of the data processing steps and dataset statistics in the Appendix Section E.*

### C. Did you run computational experiments?

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

*The detailed experimental setup is reported in Appendix Section B, where we explain a subset of the parameters; the remaining ones are chosen based on empirical knowledge.*

- 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?

*We run all evaluations with temperature fixed at 0 under the same online setting, as documented in the Appendix Section B. This setting makes the outputs of an identical model deterministic, so repeated evaluations yield identical results. Therefore, descriptive statistics such as error bars are not applicable.*

---

*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.?

*Not applicable. This work does not involve human subjects or 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)?

*Not applicable. This work does not involve human subjects or annotators.*

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)?

*Not applicable. This work does not involve human subjects or annotators.*

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

*Not applicable. This work does not involve human subjects or annotators.*

**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?

*We only employed AI tools to refine the code and polish the language in our paper and checklist, without influencing the design of experiments, the analysis, or the reported results.*