

Responsible NLP Checklist

Paper title: *DARL: Encouraging Diverse Answers for General Reasoning without Verifiers*

Authors: *Chongxuan Huang, Lei Lin, Xiaodong Shi, Wenping Hu, Ruiming Tang*

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?
Section Limitations

B. Did you use or create scientific artifacts? (e.g. code, datasets, models)

N/A 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?
This work exclusively uses existing open-source benchmark datasets that are publicly available and widely used in prior research. The data were not newly collected by the authors, and no additional personally identifying information or offensive content was introduced or processed beyond the original datasets.

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 5

C. Did you run computational experiments?

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

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 6

D. Did you use human annotators (e.g., crowdworkers) or research with human subjects?

N/A 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.?
This work does not involve human participants or formal human annotation procedures. The training

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.

and evaluation data are derived from existing open-source datasets. We only performed limited manual inspection by the authors to verify the correctness of a small subset of synthetically generated alternative answers, rather than providing structured instructions or conducting human subject studies.

- 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 study does not involve human participants, crowdsourced workers, or paid annotators. All experiments are conducted using existing open-source datasets, and no recruitment or compensation procedures are applicable.

- 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 exclusively uses existing open-source datasets that were released for research purposes under their original licenses. We do not collect new data or interact with data subjects, and consent procedures were handled by the original dataset creators.

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

Our work uses only publicly available datasets released for research purposes. No new data collection involving human subjects was conducted, so no ethics review board approval was required.

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 used AI assistants solely for tasks such as typo correction and minor language polishing. The AI did not contribute to the scientific content, analysis, or creation of the paper. All research ideas, methodology, experiments, and results were generated and written by the authors.