Responsible NLP Checklist

Paper title: *Simplify-Pro: A Two-level and Progressive LLM-based Framework for Auto Long Text Simplification*
Authors: *Peng Zhou, Guangxin Li, Xiaoying Huang, Yiming 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?
Our work focuses on improving automatic long text simplification from a technical perspective, aiming to enhance readability and accessibility while preserving the original meaning of the input text. The proposed framework does not introduce new data sources, user profiling, or decision-making mechanisms that could affect individuals or groups. It operates solely on textual inputs provided at inference time and does not generate, modify, or infer sensitive personal information. Moreover, the method does not involve user interaction, deployment in safety-critical scenarios, or automated actions that could cause real-world harm. As a result, we do not discuss specific ethical, social, or safety 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?
Our work does not involve the collection of new data or the creation of datasets containing personal or user-generated content. All datasets used in our experiments are publicly available benchmarks that have been previously released. Any necessary anonymization or filtering of personal or sensitive information was handled in the original dataset construction process by their original authors.

☑ 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

☑ C. Did you run computational experiments?

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

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.

☑ 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 4.2, Section 4.3, Section 4.4, Section 4.5, Section4.6, Appendix F, Appendix G, Appendix I, Appendix J

☑ **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.?
Section4.6

☒ 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)?
One of the authors, a professional English teacher, invited several researchers to participate in the annotation. Participation was voluntary, and all annotators provided informed consent. The annotation task was lightweight and was conducted as part of academic collaboration.

☒ 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)?
All datasets used in this work are publicly available benchmark datasets released by their original authors, who are responsible for obtaining any necessary consent. We use these datasets in accordance with their licenses and intended research purposes. In addition, our human evaluation involves voluntary participation with informed consent, and does not involve the collection of sensitive personal data.

☒ D4. Was the data collection protocol approved (or determined exempt) by an ethics review board?
The study does not involve the collection of sensitive personal data. All datasets used are publicly available and were collected and released by their original authors, who are responsible for any required ethics approvals. Human evaluation was conducted with voluntary participation and informed consent, and is considered none-risk, which typically does not require formal ethics 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?
AI assistants were not used to generate scientific content, experimental results, or research claims in this work. All core ideas, methods, experiments, and analyses were produced by the authors.