

Responsible NLP Checklist

Paper title: *Beyond Hard Masks: Progressive Token Evolution for Diffusion Language Models*
Authors: *Linhao Zhong, Linyu Wu, Bozhen Fang, Tianjian Feng, Chenchen Jing, Wen Wang, Jiaheng Zhang, Hao Chen, Chunhua Shen*

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 methodology relies exclusively on publicly available, pre-trained open-source models and established training datasets. These foundational components have already undergone extensive rigorous peer review and safety audits by the global research community. Furthermore, our application does not introduce novel architectural vulnerabilities or high-risk data processing, ensuring that the project remains within the safety and ethical boundaries established by the original developers and the open-source community.

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?

We use only publicly available benchmark datasets that are standard in prior research. These datasets have already been vetted by the community, and our work does not involve collecting new data, annotating personal information, or generating content that could be offensive.

- 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?

Appendix C

C. Did you run computational experiments?

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

Section 5, 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 5, Appendix C

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 do not use human 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 do not use human 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)?

Use public open source datasets with license checked.

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

Use public open source datasets with license checked.

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 clarify the role of AI assistants in preparing this work. They were used exclusively for language polishing, such as refining grammar, style, and readability.