

Responsible NLP Checklist

Paper title: *PHOTON: Hierarchical Autoregressive Modeling for Lightspeed and Memory-Efficient Language Generation*

Authors: *Yuma Ichikawa, Naoya Takagi, Takumi Nakagawa, Yuzi Kanazawa, Akira Sakai*

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?

See the Limitations section.

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?

See the Limitations section.

- 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?

See Section 3 and Appendix D.1.

C. Did you run computational experiments?

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

See Section 3 and Appendices D.1-D.2.

- 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 report point estimates for our main experimental results, but we do not include multi-seed statistics, confidence intervals, or error bars. This is because our study involves large-scale language model pretraining and long-context inference benchmarking, for which repeating the full training and evaluation pipeline across multiple random seeds would require substantial additional compute and time. We therefore prioritize a broad comparison across model architectures and efficiency settings within a feasible compute budget, and leave a more exhaustive multi-seed statistical analysis to future work.

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 did not use human participants 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)?

We did not recruit or pay human participants 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)?

We did not collect data directly from human participants for this study.

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

The study did not involve human-subject data collection requiring ethics board approval.

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 only for light language polishing and minor code revision, not for the core research, methodology, or results.