

Responsible NLP Checklist

Paper title: *Apertus: Democratizing Open and Compliant LLMs for Global Language Environments*

Authors: *Alejandro Hernandez-Cano, Alexander Hgele, Allen Hao Huang, Angelika Romanou, Antoni-Joan Solergibert, Barna Psztor, Bettina Messmer, Dhia Garbaya, Eduard Frank urech, Ido Hakimi, Juan Garcia Giraldo, Mete Ismayilzada, Negar Foroutan, Skander Moalla, Tiancheng Chen, Vinko Sabolec, Yixuan Xu, Michael Aerni, Badr AlKhamissi, Ins Altemir Marinas, Mohammad Hossein Amani, Matin Ansaripour, Ilija Badanin, Harold Benoit, Emanuela Boros, Nicholas John Browning, Fabian Bsck, Maximilian Bther, Niklas Canova, Camille Challier, Clment Charmillot, Jonathan Coles, Jan Milan Deriu, Arnout Devos, Lukas Drescher, Daniil Dzenhaliou, Maud Ehrmann, Dongyang Fan, Simin Fan, Silin Gao, Miguel Gila, Mara Grandury, Diba Hashemi, Alexander Miserlis Hoyle, Jiaming Jiang, Mark Klein, Andrei Kucharavy, Anastasiia Kucherenko, Frederike Lbeck, Roman Machacek, Theofilos Ioannis Manitaras, Andreas Marfurt, Kyle Matoba, Simon Matrenok, Henrique Mendona, Fawzi Roberto Mohamed, Syrielle Montariol, Luca Mouchel, Sven Najem-Meyer, Jingwei Ni, Gennaro Oliva, Matteo Pagliardini, Elia Palme, Andrei Panferov, Lo Paoletti, Marco Passerini, Ivan Pavlov, Auguste Poiroux, Kaustubh Ponshe, Nathan Ranchin, Javier Rando, Mathieu Sauser, Jakhongir Saydaliev, Mukhammadali Sayfiddinov, Marian Schneider, Stefano Schuppli, Marco Scialanga, Andrei Semenov, Kumar Shridhar, Raghav Singhal, Anna Sotnikova, Alexander Sternfeld, Ayush Kumar Tarun, Paul Teiletche, Jannis Vamvas, Xiaozhe Yao, Hao Zhao, Alexander Ilic, Ana Klimovic, Andreas Krause, Caglar Gulcehre, David Rosenthal, Elliott Ash, Florian Tramr, Joost VandeVondele, Livio Veraldi, Martin Rajman, Thomas C. Schulthess, Torsten Hoefler, Antoine Bosselut, Martin Jaggi, Imanol Schlag*

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?

Appendix F.5

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?

Section 2

- 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?

Appendices DF

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.

C. Did you run computational experiments?

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

Section 3

- 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?

Appendix F

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.?

(left blank)

- 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)?

(left blank)

- 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)?

Section 2

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

(left blank)

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?

Appendix E