

Responsible NLP Checklist

Paper title: *Alexandria: A Multi-Domain Dialectal Arabic Machine Translation Dataset for Culturally Inclusive and Linguistically Diverse LLMs*

Authors: Abdellah EL MEKKI, Samar M. Magdy, Houdaifa Atou, Ruwa AbuHweidi, Baraah Qawasmeh, OMER NACAR, Thikra Al-hibiri, Razan Saadie, Hamzah A. Alsayadi, Nadia GHEZAIEL HAMMOUDA, Alshima Mohammed Alkhazimi, Aya Hamod, Al-Yas Yaqoob Al-Ghafri, Wesam El-Sayed, Asila ismail al sharji, Mohamad Ballout, Anas Belfathi, karim Ghaddar, Serry Sibae, Alaa Aoun, Aeej Mohammed aseri, Lina Abureesh, Ahlam Bashiti, Majdal Yousef, Abdulaziz Hafiz, Yehdih Mohamed, Emira Hamedtou, Brakehe Emehah, Rahaf Alhamouri, Youssef Nafea, Aya El aatar, Walid Al-Dhabyani, Emhemed S. Hamed, Sara Shatnawi, Fakhraddin Alwajih, Khalid Elkhidir, Ashwag Alasmari, Abdurrahman Gerrio, Omar Said Alshahri, AbdelRahim A. Elmadany, Ismail Berrada, Amir Azad Adli Al-kathiri, Fadi zaraket, Mustafa Jarrar, Yahya MOHAMED EL HADJ, Hassan Alhuzali, Muhammad Abdul-Mageed

How to read the checklist symbols:

- the authors responded 'yes'
- the authors responded 'no'
- 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?
(left blank)

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?
A.1 English Source Quality Assurance

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?
3.4 Alexandria Statistics

C. Did you run computational experiments?

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

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?
5 Results and Discussion

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.?

3.2 Human Dialectal Arabic Translation – 3.3 Translation Revision – A.4 Annotation Platform

- 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)?

Ethics Statement

- 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)?

Ethics Statement

- 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?

AI use was limited to grammar correction and spell check as English is not our official language.