

Responsible NLP Checklist

Paper title: *ROMA: Real-time Omni-Multimodal Assistant with Interactive Streaming Understanding*

Authors: *Tian Xueyun, Wei Li, Bingbing Xu, Heng Dong, Yuanzhuo Wang, Huawei 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?

Ethical Statement (unnumbered section after Limitations): we discuss risks including unauthorized surveillance/privacy infringement, hallucinations/bias inherited from the base LLM, and the need for human oversight in critical applications.

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 used publicly available datasets consistent with their original licenses, but the paper does not include a dedicated discussion of PII/offensive-content screening or anonymization procedures.

- 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 3.2.1 and Figure 4; Appendix A.4. We report dataset statistics and task taxonomy, including 676,731 total samples with 27,207 alert samples, 108,986 narration samples, and 540,538 reactive QA samples.

C. Did you run computational experiments?

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

Section 5.1, Section 5.4, Appendix A.2, and Appendix A.5. We report the training/inference setup and key hyperparameters, including sampling rate, image resolution cap, sequence length, hardware, batch size, training steps, and sensitivity analyses for the timing-related hyperparameters.

- 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 benchmark results, ablations, and sensitivity analyses, but we do not report multi-run descriptive statistics such as mean/std, confidence intervals, or error bars; the reported numbers appear to be single-run results.

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.?
(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)?
(left blank)

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?
We use AI for grammar checking and paper polishing.