

Responsible NLP Checklist

Paper title: *ChunQiuTR: Time-Keyed Temporal Retrieval in Classical Chinese Annals*

Authors: *Yihao Wang, Zijian He, Jie Ren, Keze Wang*

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?

Limitations (p. 9; Potential Risks paragraph within 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?

No. The paper does not include a dedicated discussion of a PII/offensive-content screening or anonymization pipeline. The dataset is derived from publicly available pre-modern historical texts about historical figures rather than living individuals, so no anonymization was performed.

- 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 3.3 and Appendix A.5 (especially Table 9).

C. Did you run computational experiments?

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

Section 5.1 and Appendix B.1.1B.1.2.

- 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?

Section 5 and Appendix B.1.2. Results are reported as single-run evaluations of checkpoints selected by validation Recall@1; we do not report multi-run error bars or variance estimates.

D. Did you use human annotators (e.g., crowdworkers) or research with human subjects?

- ^{N/A} 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.?

N/A. Manual verification and correction were performed by the authors during dataset curation rather than by recruited participants or external annotators.

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.

N/A 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)?

N/A. The manual verification was conducted by the authors; no external participants or paid annotators were recruited.

N/A 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)?

N/A. The dataset is derived from publicly available pre-modern historical texts rather than data collected from living individuals or human subjects.

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

N/A. The work does not involve human-subject experimentation or data collection from living individuals; manual curation was performed by the authors on publicly available texts.

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 assistants were used in a limited supporting role for language polishing, brainstorming alternative phrasings, and minor coding assistance. All research design, dataset construction, experiments, result verification, and final manuscript revisions were performed and checked by the authors, who take full responsibility for the content of the paper.