# Responsible NLP Checklist

Paper title: *ToneCraft: Cantonese Lyrics Generation with Harmony of Tones and Pitches*
Authors: *Junyu Cheng, Chang Pan, Shuangyin Li*

> 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?
*The research does not involve sensitive data, real-world deployment, or applications that pose significant societal risks.*

☑ **B. Did you use or create scientific artifacts? (e.g. code, datasets, models)**

☑ B1. Did you cite the creators of artifacts you used?
*References*

☒ B2. Did you discuss the license or terms for use and/or distribution of any artifacts?
*All artifacts are used in accordance with their original licenses.*

☒ B3. Did you discuss if your use of existing artifact(s) was consistent with their intended use, provided that it was specified? For the artifacts you create, do you specify intended use and whether that is compatible with the original access conditions (in particular, derivatives of data accessed for research purposes should not be used outside of research contexts)?
*All artifacts were used strictly for non-commercial research purposes, consistent with their canonical academic applications as documented in their original papers.*

☒ 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?
*All datasets used are pre-processed public benchmarks with no PII or offensive content, as confirmed in their original releases.*

☒ B5. Did you provide documentation of the artifacts, e.g., coverage of domains, languages, and linguistic phenomena, demographic groups represented, etc.?
*All artifacts are widely-used benchmarks with comprehensive documentation in their original papers*
*B6 Statistics For Data: Yes*

☑ 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?
*5*

---

☑ **C. Did you run computational experiments?**

☑ C1. Did you report the number of parameters in the models used, the total computational budget (e.g., GPU hours), and computing infrastructure used?
*5*

☑ C2. Did you discuss the experimental setup, including hyperparameter search and best-found hyperparameter values?
*D.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?
*5*

☑ C4. If you used existing packages (e.g., for preprocessing, for normalization, or for evaluation, such as NLTK, SpaCy, ROUGE, etc.), did you report the implementation, model, and parameter settings used?
*F.1*

☑ **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.?
*All human participation involved minimal-risk tasks with no foreseeable harm.*

☒ 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)?
*Student volunteers participated voluntarily without compensation, following our institution's ethical guidelines.*

☒ 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)?
*All datas provide documented consent for research use.*

☒ D4. Was the data collection protocol approved (or determined exempt) by an ethics review board?
*The study used only pre-existing, anonymized public datasets where ethics approval was obtained during original data collection.*

☒ D5. Did you report the basic demographic and geographic characteristics of the annotator population that is the source of the data?
*Annotator demographics were not recorded to ensure privacy protection, as the study focused solely on technical annotation quality.*

☒ **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?
*All content is the original work of the authors.*