

## Responsible NLP Checklist

Paper title: *WebSynthesis: World Model-Guided Monte Carlo Tree Search for Efficient WebAgent Trajectory Synthesis*

Authors: *Yifei Gao, Junhong Ye, Yifan Yang, Jiaqi Wang, Yi Zhang, Zhang Ruichen, Jitao Sang*

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?

*Limitation 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?

*The paper only describes data sources/uses (e.g., public trajectory data, synthetic data) but never discusses PII/offensive content checks or protection/anonymization steps.*

- 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?

*Experimental Analysis*

### C. Did you run computational experiments?

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

*4 Experiment 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?

*3 Methodology, 4 Experiment*

### 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.?

*The paper relies on model training, simulations, and benchmark evaluations, with no participant-related experiments, so no participant instructions (e.g., text, screenshots) are reported.*

---

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.

- 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)?

*The papers research relies on model training, simulations, and public benchmarksno participant involvement exists, so no recruitment/payment info or related discussions are reported.*

- 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)?

*The paper uses public trajectory data, synthetic data, and TextUI structured data (no personal data), so it does not discuss obtaining consent from data owners or explaining data use.*

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

*The paper uses public trajectory data and generates synthetic data (no human-subject data collection), so it does not mention ethics review board approval or exemption for data collection protocols.*

**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 the research, specifically to assist with coding during the study process. However, the paper uses LLMs (e.g., GPT-4 and Qwen-max) as technical components, not AI assistants, and never includes info about AI assistant use.*