

Reproducibility Review of: "Geo-Alignment of Vague Cognitive Regions: Representing Uneven Cognitive Geographies of Large Language Models"



Reproduction Metadata

Submission Title: Geo-Alignment of Vague Cognitive Regions: Representing Uneven Cognitive Geographies of Large Language Models

Post Review Title: ---

Authors:

- Mina Karimi (Department of Geography and Regional Research, University of Vienna, Austria)
- Krzysztof Janowicz (Department of Geography and Regional Research, University of Vienna, Austria)
- Zilong Liu (Department of Geography and Regional Research, University of Vienna, Austria)
- Songlin Wang (Department of Geography and Regional Research, University of Vienna, Austria)
- Annika Süß (Department of Geography and Regional Research, University of Vienna, Austria)

Ref. Paper: <https://doi.org/10.5194/agile-giss-7-6-2026>

Reproducibility Reviewer: Franz Welscher (Department of Geoinformatics, University of Salzburg, Austria) | (<https://orcid.org/0000-0003-2432-1880>)

Date of Review Completion: 28.05.2026

Review Repository: <https://github.com/reproducible-agile/reviews-2026/tree/main/reports/031>

Paper Repository: <https://github.com/MinaKarimi/Geo-alignment-of-AI-in-VCRs>

Summary

The study is partially reproducible due to external resource constraints. While the provided code, data, and documentation (e.g., Readme.md) are clear and complete, a full reproduction was not possible because:

- **Data generation** relied on paid APIs (OpenAI, DeepSeek, Google Gemini), which were not accessible for re-running.

All necessary files and instructions to reproduce the results are provided by the authors, but full validation depends on access to the specified APIs.

Reproducibility Reviewer Notes

Followed README.md file of manuscript repository.

1. Reproduce Generated Data

The data in this research was generated by multiple LLMs (GPT-4o, Gemini-2.5-Flash, Deepseek-V3) through paid APIs (OpenAI API, DeepSeek API, Google Gemini API). Therefore the data generation was not repeated.

2. Reproduce Indices

The authors provide a notebook with the code for recalculating the indices, as well as the necessary data. The values of the tables displayed in the paper were reproduced by rerunning the notebook. Below the reproduced results for the Sahara Region are displayed.

Full Jaccard Similarity Matrix:

	B_GPT4o_En	B_DS_En
B_GPT4o_En	1.00	0.23
B_DS_En	0.23	1.00

Full Dice Similarity Matrix:

	B_GPT4o_En	B_DS_En
B_GPT4o_En	1.00	0.37
B_DS_En	0.37	1.00

Full Pearson Correlation Matrix:

	S_GPT4o_En	S_DS_En
S_GPT4o_En	1.0	0.6
S_DS_En	0.6	1.0

3. Reproduce Figures

Figures of the manuscript were created with ArcGIS. All data was provided for potential recreation in ArcGIS.

Citing this document

This report is part of the reproducibility review at the AGILE conference. For more information see <https://reproducible-agile.github.io/>. This document is published on OSF/ResearchEquals at <https://doi.org/10.17605/OSF.IO/WH6GD>.

To cite the report use

Welscher, Franz. 2026. "Reproducibility Review Of: Geo-alignment of Vague Cognitive Regions: Representing Uneven Cognitive Geographies of Large Language Models." OSF. June 13. doi:10.17605/OSF.IO/WH6GD.