CartoAgent: a multimodal large language model-powered multi-agent cartographic framework for map style transfer and evaluation

Chenglong Wang[ab], Yuhao Kang[c], Zhaoya Gong[ab#], Pengjun Zhao[ab]

Yu Feng[d], Wenjia Zhang[ab] and Ge Li[e]

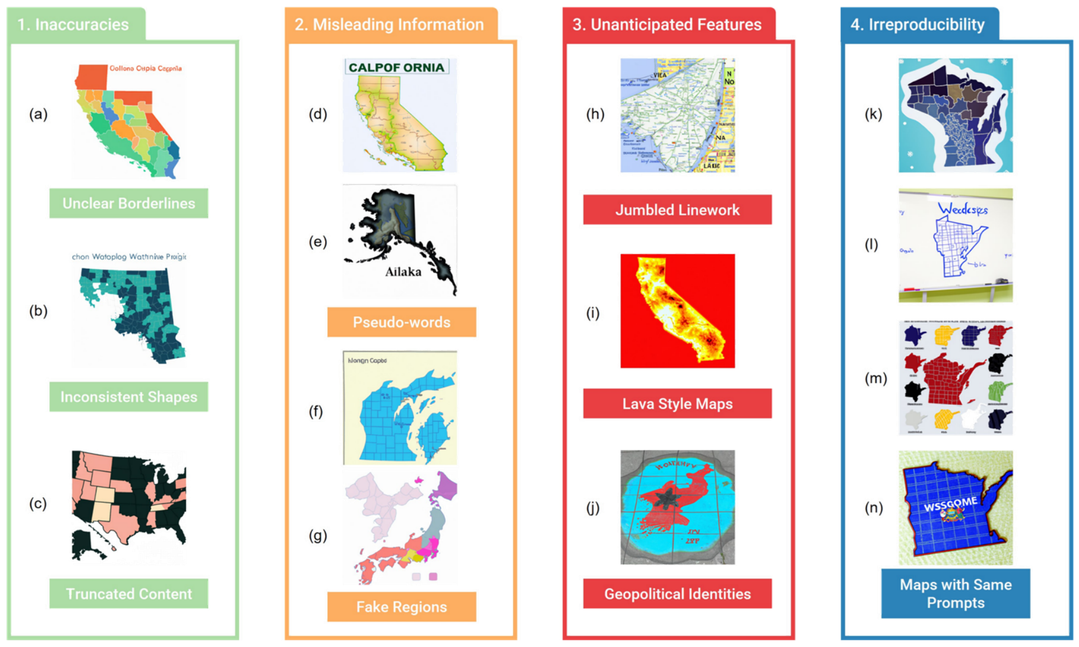
[a] School of Urban Planning and Design, Peking University Shenzhen Graduate School, Shenzhen, China;
[b] Key Laboratory of Earth Surface System and Human-Earth Relations of Ministry of Natural Resources of China, Peking University Shenzhen Graduate School, Shenzhen, China;
[c] Department of Geography and the Environment, University of Texas at Austin, Austin, USA;
[d] School of Engineering and Design, Technical University of Munich, Munich, Germany;
[e] School of Electronic and Computer Engineering, Peking University Shenzhen Graduate School, Shenzhen, China;
[#] Corresponding author: Zhaoya Gong (z.gong@pku.edu.cn)

[Introduction]

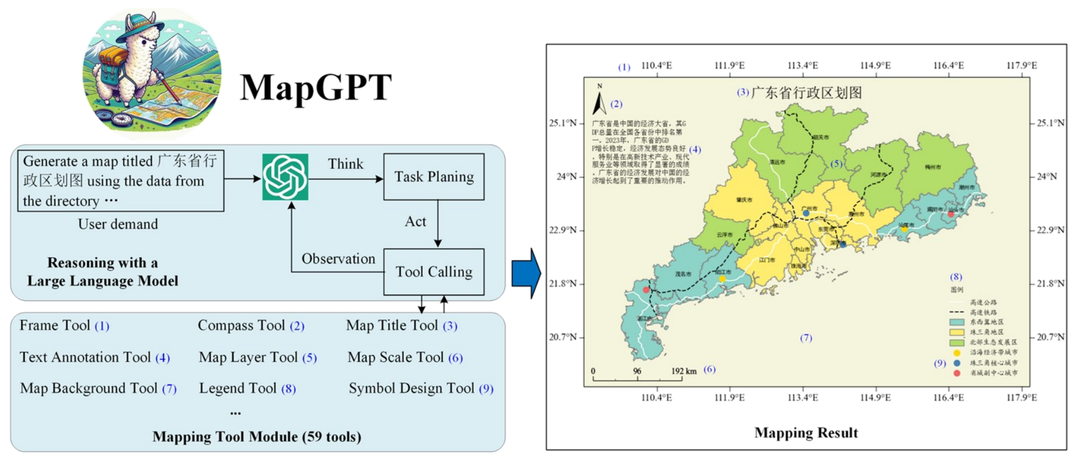
How can Generative AI (GenAI) benefit the map-making process?

- When utilizing **text-to-image models** like DALL·E and Midjourney to generate maps, the resulting raster-based map images often contain **inaccurate and unreliable geographic information**.
- When leveraging **text-to-text models** like ChatGPT, researchers could successfully simulate cartographers' actions, significantly automating the map-making process. However, this method **cannot handle the artistic side of maps**.

(A)



(B)



(A) Example AI-generated maps: (1) inaccuracies, (2) misleading information, (3) unanticipated features, and (4) irreproducibility (Kang et al., 2023). (B) The framework of MapGPT. Users first describe their demands by natural language text; subsequently, a large language model (LLM) is used to interpret demands, plan tasks, and call appropriate tools in the Mapping Tool Module for different mapping tasks (Zhang et al., 2024).

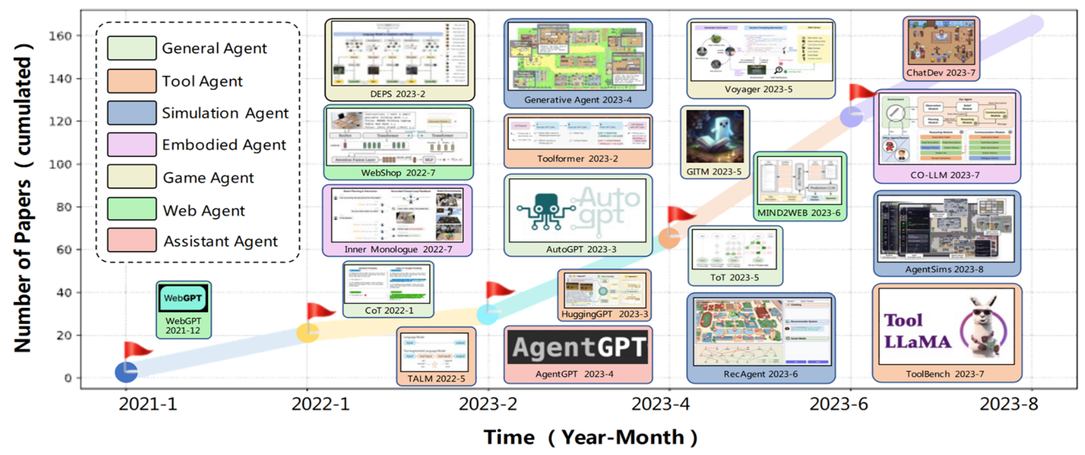
Is it possible to develop a novel GenAI model that can efficiently generate visually appealing maps while maintaining accurate geographic information?

- **Art of cartography:** proper and appealing color schemes and symbols.
- **Science of cartography:** effective communication of the content on maps, through an objective process of establishing and following laws/rules for map design.

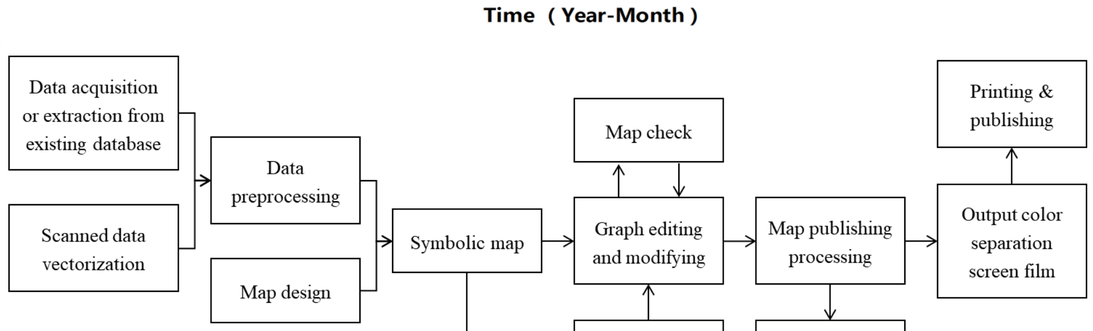
[Motivations]

- Multimodal large language models (**vision-language ability**)
- LLM-powered autonomous agents (**task management ability**)
- Operation process of digital mapping

(A)

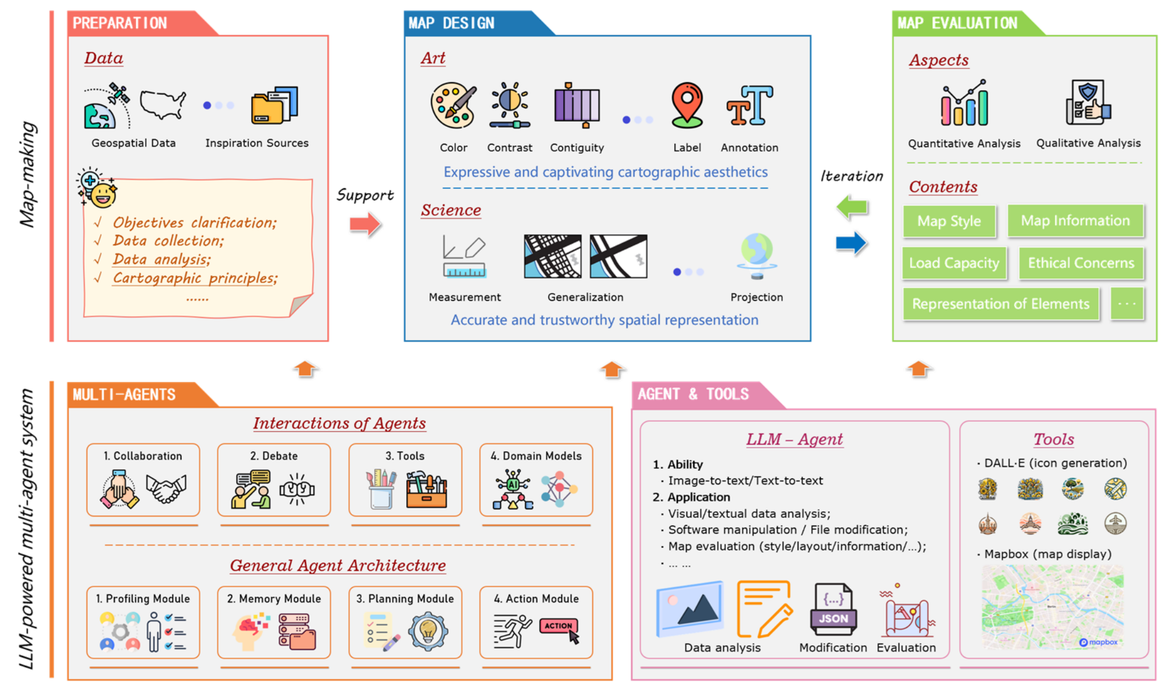


(B)



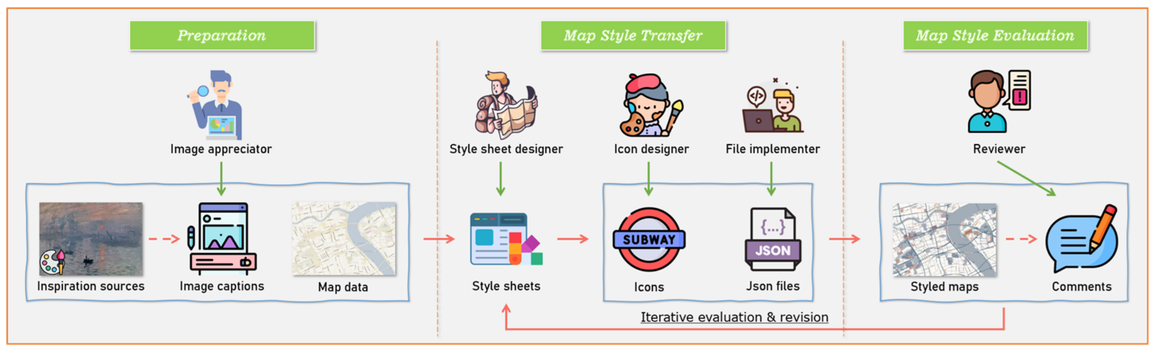
(A) Illustration of the growth trend in the field of LLM-based autonomous agents (Wang et al., 2023). (B) The operation process of digital mapping (Wang & Wu, 2021).

[Framework]



CartoAgent: a conceptual cartographic framework powered by multimodal large language models.

[Applications]

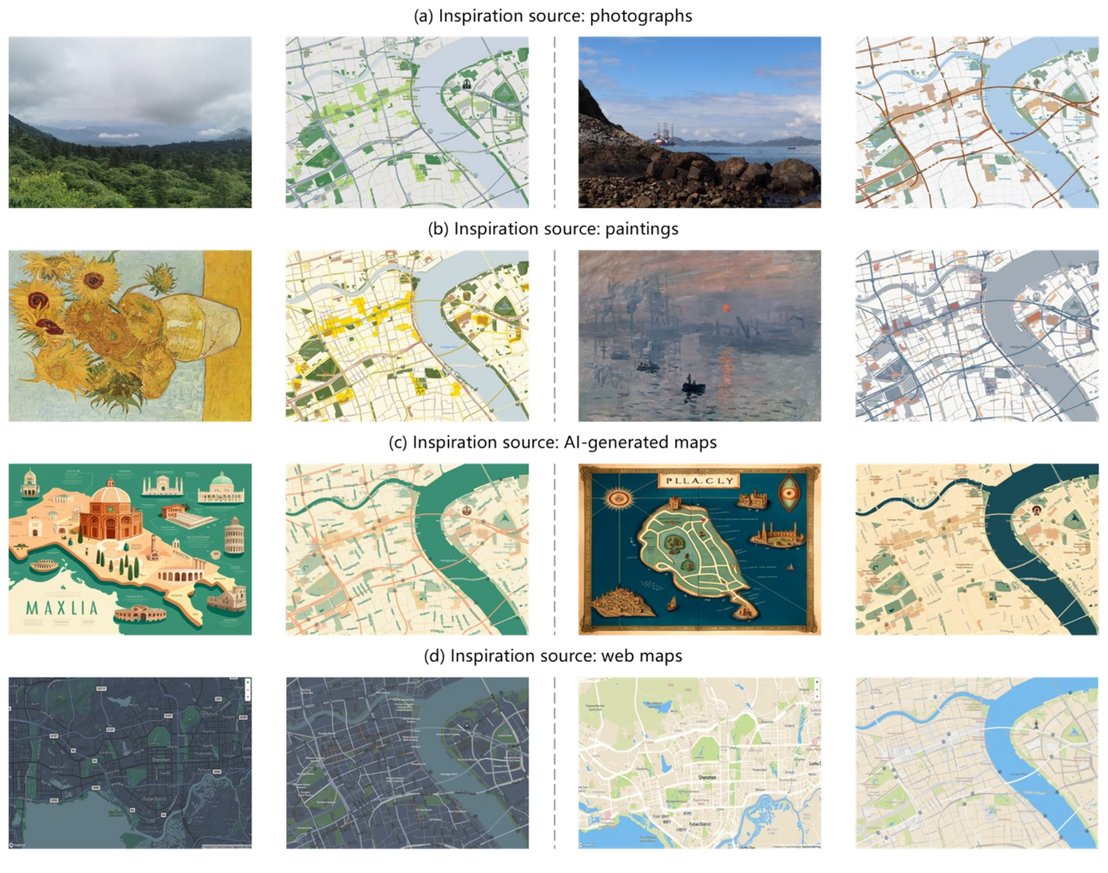


Map style transfer and evaluation:

- The process of **reproducing artistic styles from existing maps, paintings, or other visual artwork into input map data**.
- **Role assignment:** image appreciator, style sheet designer, icon designer, file implementer & reviewer.

[Results & Discussions]

Neighborhood-level (with evaluation)



Discussions

- What knowledge is required for map style transfer?
- How to evaluate the results of map style transfer?
- How to build a robust cartographic system for map-making?
- Ethical issues related to map style transfer.

[References]

Kang, Y., Zhang, Q., & Roth, R. (2023). The ethics of ai-generated maps: A study of dalle 2 and implications for cartography. arXiv preprint arXiv:2304.10743.

Zhang, Y., He, Z., Li, J., Lin, J., Guan, Q., & Yu, W. (2024). MapGPT: An Autonomous Framework for Mapping by Integrating Large Language Model and Cartographic Tools.

Wang, L., Ma, C., Feng, X., Zhang, Z., Yang, H., Zhang, J., ... & Wen, J. R. (2023). A survey on large language model based autonomous agents. arXiv preprint arXiv:2308.11432.

Wang, J., & Wu, F. (2021). Advances in Cartography and Geographic Information Engineering. Singapore: Springer.