International Cartographic Association
Association Cartographique Internationale

CartoDirector: A geospatial storyteller that converts text into dynamic cartographic transformations using LLMs

Pengbo Li [a]*, Haowen Yan [a]*, Jiannan Li [b], Xiaomin Lu[a], Weifang Yang [a]

[a] *Faculty of Geomatics, Lanzhou Jiaotong University, Pengbo Li - pengbo.li@lzjtu.edu.cn, Haowen Yan - yanhw@mail.lzjtu.cn, Xiaomin Lu - xiaominlu08@mail.lzjtu.cn, Weifang Yang- ywfo208@lzjtu.edu.cn*
[b] *School of Computing and Information Systems, Singapore Management University, Jiannan Li - jiannanli@smu.edu.sg*

* Corresponding author

Keywords: cartographic transformation, visual storytelling, large language model

Abstract:

Maps transcend cultural and linguistic boundaries to communicate spatial and temporal information. People are used to communicating through natural language, which is descriptive but offers less intuitive spatial information. Cartographic-based visual storytelling makes spatial data more intuitive and accessible (Roth 2020). Therefore, we proposed a novel autonomous mapping system, CartoDirector, which leverages large language models (LLMs) to transform textual narratives into dynamic cartographic processes. This system bridges the gap between storytelling and geographic visualization, enabling users to generate spatially and temporally coherent maps from descriptive texts. Just as a film director interprets a script by deciding on elements such as camera angles and lighting, CartoDirector transforms a narrative into maps. By understanding the context and spatial relationships described in the text, the system generates spatially coherent maps that reflect the story's geographic and temporal aspects. By harnessing the advanced natural language understanding capabilities of LLMs, CartoDirector enriches input texts, interprets contextual details, identifies geographic references, and converts them into appropriate cartographic processes.

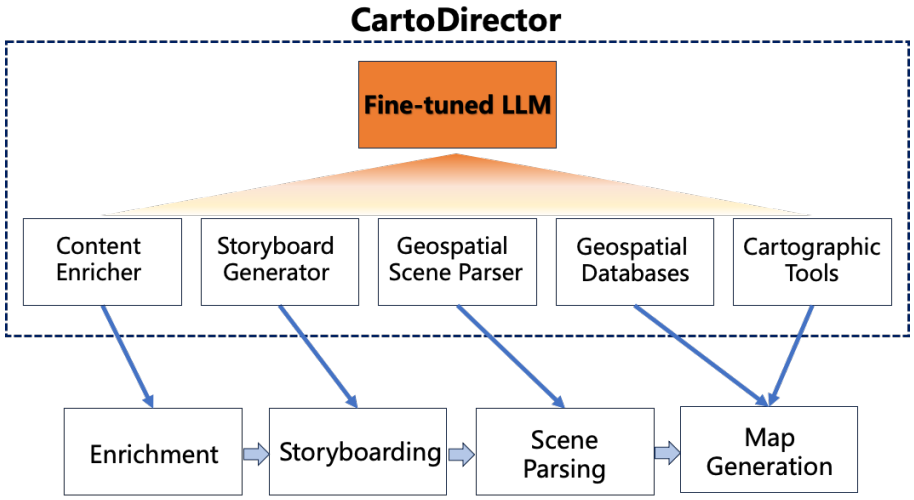

Figure 1. Framework of CartoDirector

Many researchers have studied large language model-based geospatial applications. For example, MapGPT is an intelligent mapping framework that interacts with natural language descriptions (Zhang et al. 2024), and GeoTool-GPT is a trainable method to enable LLMs to master GIS tools (Wei et al. 2024). CartoDirector distinguishes itself by not just focusing on generating maps or calling GIS tools, but on understanding a story and transforming it into a visual storytelling plan using cartographic transformations, including anchor translation, scaling, perspective changes, style transfer and modality transfer. As shown in Figure 1, CartoDirector consists of a fine-tuned LLM and five key modules: content enricher, storyboard generator, geospatial scene parser, geospatial databases and cartographic tools. Each module is under the guidance of the fine-tuned LLM. CartoDirector follows a structured four-step workflow:

1) Enrichment: The content enricher accepts a story text and enriches it to an extended text with detailed mapping information.

2) Storyboarding: The storyboard generator can automatically cut the extended text into sequential text scripts, setting the stage for how the story will unfold visually.

3) Scene parsing: Subsequently, the geospatial scene parser can convert each script to an object-oriented scene file with cartographic information about geographic entities, scale, style, perspective, modality, etc, critical for creating expressive map representations.

4) Map generation: Finally, appropriate geospatial databases and cartographic tools are selected according to the scene description file, allowing for the generating of a series of maps that effectively convey the story.

Given a story text, CartoDirector can convert it into coherent maps with the proper scale, symbol, view and modality. Figure 2 illustrates the workflow of the whole process. These four steps enable the conversion of a simple story into a visually dynamic, geographically informed map narrative. The result would be a sequence of maps in an intuitive way. In a broader context, CartoDirector exemplifies the power of LLM-based generative AI (Wang et al 2024), where large language models are used not only for text generation but also for understanding and transforming a story into spatially and temporally coherent and accurate map representations.

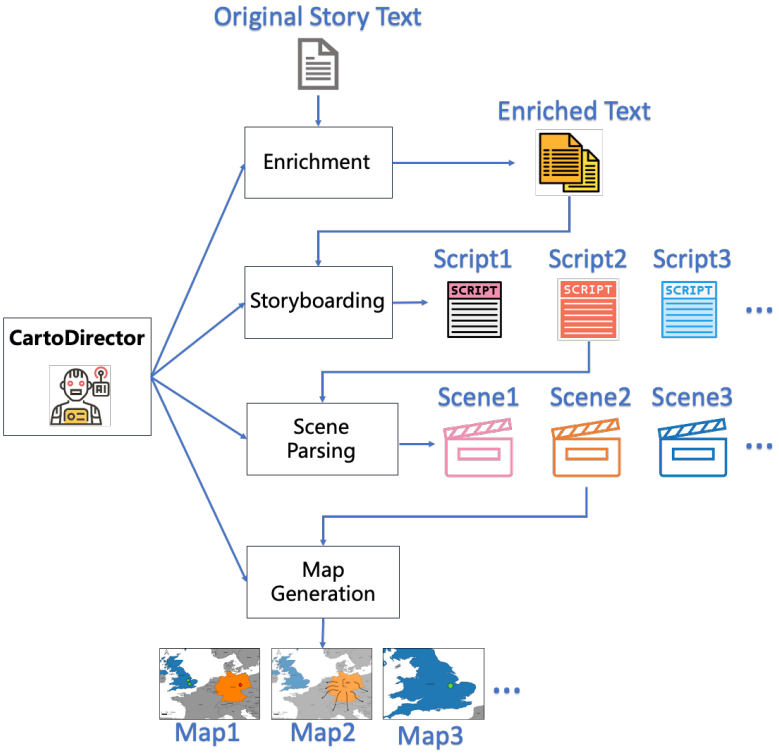

Figure 2. Workflow of CartoDirector

Acknowledgements

The authors would like to acknowledge the support of the National Natural Science Foundation of China [42430108, 42371463, 42161066], Science-Technology Foundation for Young Scientists of Gansu Province [25JRRA202]

References

Roth, R. E. (2020). Cartographic design as visual storytelling: synthesis and review of map-based narratives, genres, and tropes. The Cartographic Journal, 58(1), 83–114.

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. Cartography and Geographic Information Science, 51(6), 717–743.

Wei, C., Zhang, Y., Zhao, X., et al. (2024). GeoTool-GPT: a trainable method for facilitating Large Language Models to master GIS tools. International Journal of Geographical Information Science, 1–25.

Wang, S., Hu, T., Xiao, H., et al. (2024). GPT, large language models (LLMs) and generative artificial intelligence (GAI) models in geospatial science: a systematic review. International Journal of Digital Earth, 17(1).