

Fig. 1. Creation Examples of GeoAuthor. Among them, B, E, F, and H require data importing for creation.

APPENDIX

A. Creation Examples of GeoAuthor

In addition to the figures in the paper, we also provide additional examples (Figure 1) that showcase the text and visualizations created by GeoAuthor.

B. Preliminary Study

1) *Research Process*: We conducted research on 80 geographic articles from various sources such as The New York Times and RedNote to summarize the connection between geographic articles and geographic visualization. Our analysis process involved two main steps: First, we broke down the final presentation of geographic visualizations into components, allowing us to identify the key elements that make up these visualizations. Second, we extracted text fragments from the articles that were related to these constituent components and analyzed how these text fragments influenced or corresponded to the visual elements.

2) *Results*: We decompose the final presentation of illustrations collected from the articles into two primary layers: map background and graphics drawn on the map. For the graphics drawn on the map, we further categorized these elements based on their drawing methods. The first category includes graphics positioned by precise latitude and longitude coordinates, which are the basic elements of geographic visualization: **point**, **line**, and **area**. The second category comprises the **annotation** derived from these primary elements, such as text labels, bounding boxes, and similar elements. The following provides a brief introduction to the first category:

- Point refers to point-based visualization, which includes the most common dots, markers, and other icons of various shapes. Inserting statistical charts at different positions on a map is also a special form of it.
- Line refers to line-based visualization, which can be divided into two categories based on its drawing method: a line drawn using key points and a line drawn directly without key points. The former can include trajectories or

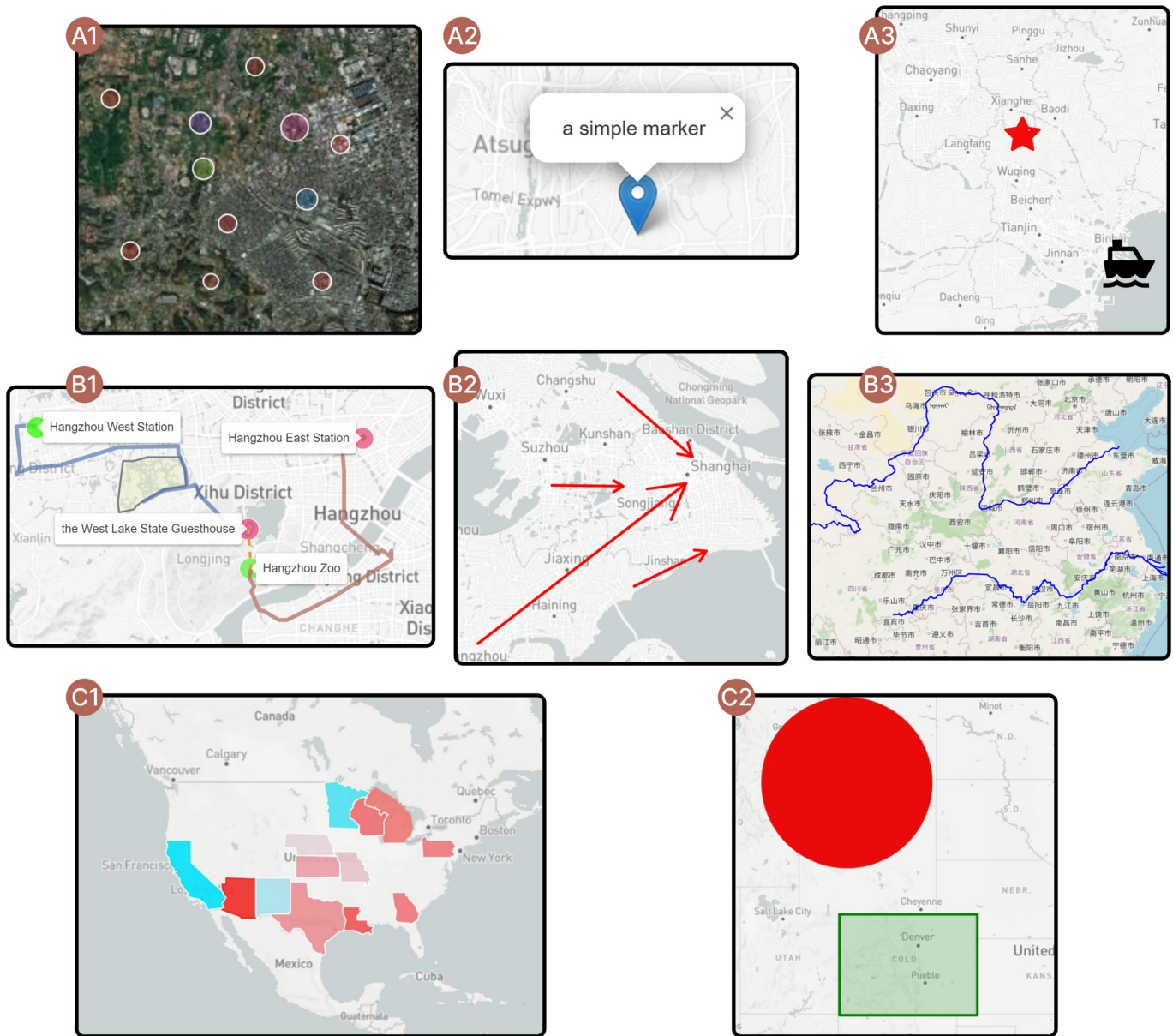


Fig. 2. The eight common visualization types we identified. (A) POINT types include three subtypes: (A1) dot, (A2) marker, and (A3) other icons. (B) LINE types include three subtypes: (B1) trajectory, (B2) direction, and (B3) line. (C) AREA types include three subtypes: (C1) irregular region and (C2) regular shape.

long directional arrows with clear start and end positions, while the latter can include national borders, roads, etc.

- Area refers to area-based visualization, which includes regular simple shapes (e.g., circles, polygons) and irregular closed regions.

Based on these findings, we further subdivided point, line, and area in the first category, summarizing a total of eight distinct visualization types: dots, markers, other icons, trajectories, directions, lines, irregular regions, and regular shapes. We believe these are the most common visualizations in geographical articles; therefore, they are adopted as the primary visualization types supported by GeoAuthor. Figure 2 presents simple examples of these eight visualization types.

C. Text-to-Visualization Generation

Section V.A of the paper introduces the generation from text to geographic visualization, where LLM prompt engineering is utilized in the process of converting text to JSON. Figure 3 shows an example of a JSON structure, and Figure 4 presents a prompt template for LINE-type JSON generation.

D. System Evaluation

The construction of our test benchmark follows a semi-automated approach using an interface we developed (Figure 5). Through this interface, users first specify the type and requirements for the text to be generated, including the visualization types that should be included and other custom requirements. The Qwen2.5-7B-Instruct model is then



Fig. 3. A structural example of JSON grammar.

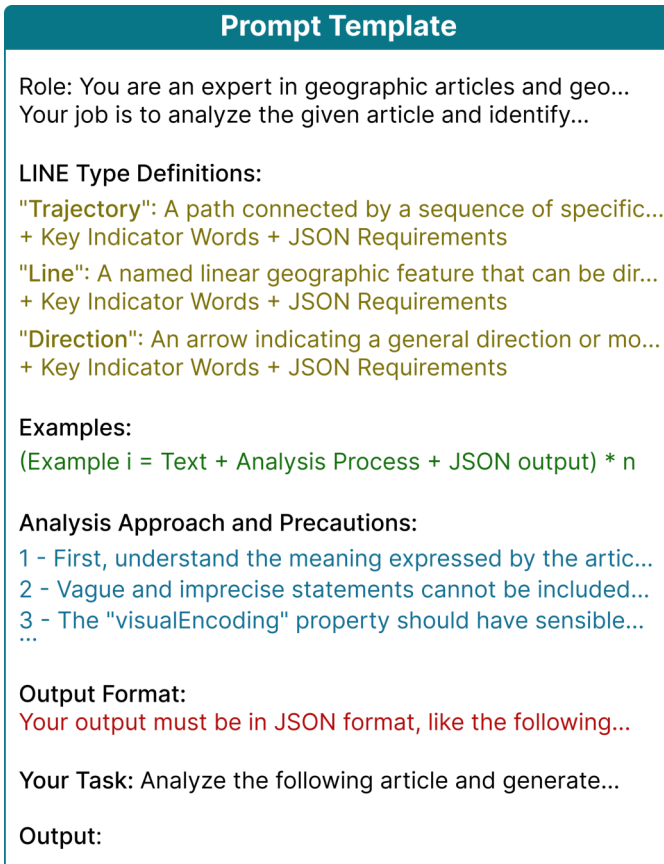


Fig. 4. A prompt template for LINE-type JSON generation.

employed to generate geographic text cases based on these specifications. Subsequently, corresponding JSON results can be generated from the geographic text by directly reusing GeoAuthor's codebase, enabling users to manually assess GeoAuthor's generation accuracy. Finally, users can adjust and modify the generated JSON to establish the ground truth. Through this workflow, once the benchmark construction is completed, we obtain the text-to-visualization accuracy of GeoAuthor.

E. User Study

1) *SUS Questionnaire*: For GeoAuthor, the raw results of the SUS questionnaire are shown in Table I.

TABLE I
RAW RESULTS OF THE SUS QUESTIONNAIRE FOR GEOAUTHOR

Participants	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10
P1	4	2	4	2	4	3	5	2	4	1
P2	3	4	3	3	4	2	3	3	3	3
P3	5	2	5	1	4	1	5	1	5	1
P4	5	2	5	1	5	1	5	1	5	2
P5	5	2	5	1	4	2	5	1	5	1
P6	5	1	5	2	5	2	4	1	5	2
P7	4	3	5	2	4	3	4	2	5	2
P8	5	2	4	2	5	2	4	2	5	1
P9	5	3	5	1	5	1	5	1	5	1
P10	5	1	5	1	5	1	5	1	5	1
P11	4	2	5	1	4	1	4	1	5	1
P12	5	1	4	1	5	1	4	1	4	1

According to the SUS standard guidelines, for even-numbered questions, lower scores are better, while for odd-numbered questions, higher scores are better. Therefore, we converted the scores for even-numbered questions from $x \in [1, 5]$ to $5-x$, and the scores for odd-numbered questions from $x \in [1, 5]$ to $x-1$. The processed results are shown in Table II.

TABLE II
PROCESSED RESULTS OF THE SUS QUESTIONNAIRE FOR GEOAUTHOR

Participants	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10
P1	3	3	3	3	3	2	4	3	3	4
P2	2	1	2	2	3	3	2	2	2	2
P3	4	3	4	4	3	4	4	4	4	4
P4	4	3	4	4	4	4	4	4	4	3
P5	4	3	4	4	3	3	4	4	4	4
P6	4	4	4	3	4	3	3	4	4	3
P7	3	2	4	3	3	2	3	3	4	3
P8	4	3	3	3	4	3	3	3	4	4
P9	4	2	4	4	4	4	4	4	4	4
P10	4	4	4	4	4	4	4	4	4	4
P11	3	3	4	4	3	4	3	4	4	4
P12	4	4	3	4	4	4	3	4	3	4

Then, we multiplied the sum of all scores by 2.5, resulting in a final SUS score of 86.67. Based on research on SUS score factors, we also calculated detailed usability and learnability scores. The usability score was 86.2 (sum of Q1-3 and Q5-9 scores, multiplied by 3.125), and the learnability score was 88.54 (sum of Q4 and Q10 scores, multiplied by 12.5).


The SUS Questionnaire for the baseline follows the same processing approach, with the processed results shown in Table III:

Select Visualization Types:

Point - Dot Point - Marker Point - Other Icon Line - Trajectory Line - Line Line - Direction

Area - Irregular Region Area - Regular Shape

Additional Requirements (Optional):

 You can specify content style, geographic region, theme, or any other preferences for the generated text.

Geographic Text:

Generated geographic text will be displayed here...

Corresponding JSON:

JSON Result Validation:

Absolutely Correct Correct but Containing Excess Elements Missing Elements Elemental Chaos

Fig. 5. The interface we developed for benchmark construction and GeoAuthor’s performance test.

TABLE III
PROCESSED RESULTS OF THE SUS QUESTIONNAIRE FOR THE BASELINE

Participants	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10
P1	1	3	3	3	1	2	3	3	3	4
P2	3	3	3	3	3	3	3	3	3	3
P3	2	4	4	4	1	4	4	4	4	4
P4	3	2	2	3	2	4	2	3	4	2
P5	2	4	3	4	2	3	4	3	4	4
P6	0	3	3	3	0	3	3	1	3	3
P7	2	1	1	1	2	1	2	2	3	3
P8	3	4	1	2	0	2	4	0	3	4
P9	0	4	4	4	0	4	0	1	4	4
P10	3	4	4	4	2	4	4	3	4	4
P11	3	4	4	4	2	4	4	3	4	4
P12	3	4	4	4	2	4	4	4	4	4

2) *System Effectiveness Questionnaire*: For the results of the custom system effectiveness questionnaire, we used Student's t-test for analysis. The detailed analysis results are shown in Table IV.

TABLE IV
ANALYSIS RESULTS FOR THE SYSTEM EFFECTIVENESS QUESTIONNAIRE

Question	MD	t(11)	p	Cohen's d	95% CI	Significant
Q1	1.083333	4.168193	0.001567	1.203254	[0.511, 1.655]	Yes
Q2	0.750000	3.446738	0.005459	0.994987	[0.271, 1.229]	Yes
Q3	1.833333	4.749665	0.000600	1.371110	[0.984, 2.683]	Yes
Q4	1.666667	4.432026	0.001008	1.279416	[0.839, 2.494]	Yes
Q5	0.500000	2.171241	0.052663	0.626783	[-0.007, 1.007]	No
Q6	0.500000	3.316625	0.006872	0.957427	[0.168, 0.832]	Yes
Q7	0.250000	1.393261	0.191054	0.402200	[-0.145, 0.645]	No
Q8	0.166667	1.483240	0.166087	0.428174	[-0.081, 0.414]	No
Q9	0.583333	2.548328	0.027076	0.735639	[0.080, 1.087]	Yes
Q10	0.333333	2.345208	0.038814	0.677003	[0.020, 0.646]	Yes