

GeoAuthor: Linking Text and Visualization for Geographic Article Authoring

Zhenning Chen, Hanbei Zhan, Shifu Chen, Zikun Deng, Di Weng, Yingcai Wu, *Senior Member, IEEE*

Abstract—Articles containing geographic information are widely distributed and commonly used in daily life, frequently incorporating geographic visualizations as illustrations. However, the creation of such articles remains cumbersome, necessitating authors to switch between authoring text and illustrations, thereby disrupting immersive writing. Our interviews corroborated this observation and revealed the primary challenge in the traditional process stems from the low synchronization frequency between text and geographic visualizations during creation, coupled with weak visual links, forcing users to mentally maintain this synchronization and thereby increasing their cognitive burden. In response, we developed GeoAuthor, which facilitates the interactive creation of geographic articles by automatically synchronizing text creation with geographic visualizations with rich visual links. This bidirectional approach ensures that the written content and visual representations remain consistent and mutually informative throughout the creation process. Our evaluation demonstrated the efficacy of GeoAuthor, indicating its capacity to streamline the process of creating geographic articles.

Index Terms—Text-Visualization Link, Geographic Visualization, Geographic Article Authoring.

I. INTRODUCTION

Articles incorporating relevant visualizations generally exhibit better retention and can enhance comprehension of illustrated content compared to purely textual content [1], [2]. This principle holds true for geographic articles as well, which we define in this paper as articles containing geographic information and illustrations. Such articles are widely distributed and commonly used in our daily lives. A prime example is found in major newspapers like The New York Times, which frequently publish news articles rich in geographic content on topics such as geopolitics, military affairs, and social issues. These articles are typically accompanied by comprehensible geographic visualizations integrated within the text, enhancing reader understanding and engagement.

However, creating geographic visualizations for geographic articles is not an easy task for article authors, primarily due to two main challenges. Firstly, many libraries and GIS (Geographic Information System) software capable of producing geographic visualizations, such as Leaflet [3] and ArcGIS [4], often require users to have coding skills or invest

significant time in learning. Most authors lack the necessary programming expertise or the time to master these specialized tools, which creates a barrier to entry or a steep learning curve. Furthermore, even if authors become proficient with these software tools, there remains a fundamental issue: these tools are primarily designed to serve geographic data analysis rather than to support textual content directly. They excel at geographic computing, data analysis, and visualization, but fall short in integrating with the narrative flow of an article. This misalignment forces authors to frequently switch between writing and illustration tasks, disrupting their writing flow and hindering immersive content creation.

The root cause of the above challenges can be attributed to an issue: there is a critical bidirectional link missing between textual authoring and geographic visualization creation. Existing work on linking text and geographic visualization either introduces only a weak linkage, limited in matching user-selected text to pre-prepared Vega-Lite charts [5], and therefore does not effectively support article authoring discussed in this paper, or does not involve generating geographic visualization from text at all [6]. Furthermore, existing studies focusing on generating geographic visualizations from text employ methods that are not directly applicable to generating illustrations for geographic articles. Traditional approaches simply extract place names from text and place labels on a map. More recent methods [7]–[10] fail to adapt to geographic data articles or the distinctive illustrations typically used in such articles, and they are designed from the reader’s perspective, which differs fundamentally from the purpose and methodology of authoring tools. These reader-oriented approaches focus on post-hoc visualization generation to help audiences understand completed articles. In contrast, authors need real-time, iterative visualization creation that integrates with their writing process and supports their storytelling intentions, rather than simply extracting geographic entities from finished text.

Thus, this work explores utilizing connections between text and geography for text-oriented geographic visualizations and convenient article authoring. We first conducted formative interviews with 10 individuals from different fields experienced in geographic article writing to understand article authors’ workflows, pain points, and desired tools when creating such articles. Based on these, we developed a research prototype GeoAuthor, which assists authors in creating geographic articles more easily and conveniently. Finally, we validated the effectiveness and usability of GeoAuthor through system evaluation and a user study.

GeoAuthor utilizes large language models (LLMs) to con-

Z. Chen, S. Chen and D. Weng are with School of Software Technology, Zhejiang University. E-mail: {czn, shifu.chen, dweng}@zju.edu.cn. D. Weng is the corresponding author.

H. Zhan and Y. Wu are with State Key Lab of CAD&CG, Zhejiang University. E-mail: {zhanhanbeiii, ycwu}@zju.edu.cn.

Z. Deng is with School of Software Engineering, South China University of Technology. E-mail: zkdeng@scut.edu.cn.

Manuscript received April 19, 2021; revised August 16, 2021.

vert author-written text into a formal grammar and then leverages it to generate corresponding geographic visualizations. This generation process creates an inherent link between the text and the geographic visualizations, establishing a dynamic relationship and dynamic interactions between the textual content and its corresponding geographic representations. With GeoAuthor, authors can focus on textual creation while seamlessly integrating geographic visualization into the writing process. This integrated approach allows article authors to maintain their writing flow while simultaneously creating and refining geographic visualizations that complement their text. The articles and illustrations created by GeoAuthor meet the illustration needs of common geographic articles.

The major contributions of this work are as follows:

- Our formative interviews to understand the pain points and desired tools when creating geographic articles.
- Our system, GeoAuthor, that links text and visualization for more convenient geographic article authoring.
- Our evaluation that demonstrates GeoAuthor’s usability and effectiveness in creating geographic articles.

II. RELATED WORK

We summarized relevant research on linking text to visualizations, geographic visualization authoring, storytelling and visualization, and AI-assisted writing.

A. Linking Text to Visualizations

There exist abundant relationships between text and visualizations [11], [12]. By connecting text and visualizations, it facilitates visualization creation for authors and enhances information comprehension for readers, transforming both authoring and consumption processes. Research has employed traditional methods to establish links between text and data [13]–[15], as well as between textual information and multimedia content [16]–[20] such as graphics [21], [22] and charts [23]–[29], to support multimodal content authoring. Beyond these works, there are also studies that leverage deep learning techniques to generate various visual forms from text, including chart-related visualizations [30]–[33] and other multimedia content [34], [35], which circumvents much of the tedious rule-based analysis compared to traditional methods.

Although these text-visualization links exist, they cannot directly support geographic visualizations due to significant differences from other visualization types. These differences include geographic name ambiguity, data complexity, and the diverse, vague expression of geographic articles whose correspondence with illustrations is not always perfectly aligned. This also motivates GeoAuthor’s visualization refinement features (Section V.C). Furthermore, while a few studies (Section I third paragraph) connect text with geographic visualizations, they cannot effectively support the interactive geographic article creation scenario this paper addresses.

B. Geographic Visualization Authoring

On the one hand, there are construction tools, including applications, programming libraries, and toolkits, that can be

applied to the creation of geographic visualizations. Widely used digital map platforms such as Google Maps and Bing Maps can provide simple visualization creations. Visualization applications [36]–[38] like Mapbox [39] and various programming libraries [40] available for geospatial data visualization such as D3 and Leaflet can also create geographic visualizations. Visual toolkits are another type of tool that can be used to create geographic visualizations [41]–[45]. For example, Geovisto [41] is a toolkit enabling non-programmers to create multi-layered map widgets for geospatial data visualization. On the other hand, there are approaches utilizing grammar to generate geographic visualizations. For instance, Vega-Lite [46] is an interactive grammar that simplifies interactive data visualization creation. The Urban Toolkit [43] is a visualization framework that streamlines urban data visualization through grammar designed for urban use cases.

Although extensive research has explored geographic visualization [47]–[56] and its creation, significant gaps remain in establishing effective connections between textual content and geographic visualization. This connection is crucial as it can provide rich interactive experiences during the author’s writing process and avoid the learning curve associated with grammar-based approaches, thereby enabling more convenient generation of geographic visualizations.

C. Storytelling and Visualization

Storytelling is an important topic in the field of visualization, while visualization is also a crucial means of conveying information in storytelling. Interactive authoring tools play a crucial role in visualization storytelling by enabling dynamic narrative creation. They enable users to create dynamic narratives through various interaction paradigms, supporting both real-time story construction and the externalization of analytical processes to facilitate knowledge sharing and collaborative understanding [57]–[62]. For example, SketchStory [61] is a digital whiteboard enabling presenters to dynamically create visualizations through sketch-based gestures during live presentations. Narrative visualization techniques can transform various information sources into structured visual stories. The principle involves extracting key elements from textual content or other materials and designing effective visual communication forms, thereby enhancing audience understanding and engagement [63]–[66]. For example, Contextifier [63] applies narrative visualization techniques to convert news articles into annotated charts. In this work, we tell geographic stories through geographic visualization and its linkage with text.

D. AI-Assisted Writing

There have been many AI-based tools and interfaces designed to assist authors in writing [67]–[70]. For example, Say Anything [67] is an interactive storytelling system that automatically provides writing suggestions for the next sentence. Different from the above studies that focus on the novelty in interaction, there are also studies emphasizing innovations in models [71]–[73]. For instance, Ammanabrolu et al. [74] present an innovative ensemble-based model for transforming plot events into coherent sentences, making the generated

stories more reasonable. While our research also employs LLMs to aid in writing, it distinctively leverages geographic information to support textual composition.

III. FORMATIVE INTERVIEWS

In this section, we present the formative interviews¹ we conducted to gain a better understanding of the workflow, pain points, and expected tools for authoring geographic articles.

A. Background and Setup

Background. Prior to our formative interviews, we conducted brainstorming sessions exploring potential methods for facilitating geographic article creation. Since we observed strong correlations between textual content and the accompanying illustrations in our collected geographic articles, we conceptualized an automatic text-based illustration generation approach. However, to avoid biasing the participants, we did not disclose this idea during the subsequent interviews.

Participants. Our 10 participants (I1–I10) were experts in geographic writing or geographic visualization, including two journalism students (I1–I2), two urban visual analytics experts (I3–I4), two computer science PhD students who post travel content on social media (I5–I6), two working journalists (I7–I8), an urban planning designer (I9), and a museum studies expert (I10). There were six females and four males, aged 22–31, all with geographic article authoring experience.

Interview Procedure. We conducted face-to-face or remote interviews lasting 50–70 minutes. Participants described their frequency of working with geographic articles, workflows for writing and creating illustrations, pain points, and desired tools. Finally, we provided New York Times text excerpts and asked participants to describe their typical approach to creating geographic visualizations for such content.

Analysis. All interviews were recorded, transcribed verbatim, and analyzed using thematic analysis [75]. We used both inductive and deductive methods to summarize preliminary codes and themes, then established the final themes through comprehensive discussion, refinement, and categorization.

B. Findings and Requirements

Regarding the writing process for paragraphs corresponding to illustrations, users behave differently, adapting to different scenarios and personal habits. The main difference lies in whether they write the textual content before or after creating the illustrations. However, regardless of the approach, all participants indicated that the process is mentally synchronized. When creating illustrations, they also conceptualize the text in their minds, and while writing the text, they simultaneously visualize the illustrations in their minds. Therefore, for the desired tool, they naturally proposed the concept of generation from text to illustrations.

As for the tools currently used for creating illustrations, participants usually struggled to find and effectively use suitable tools when authoring different geographic articles. Most

participants (I1–I2, I5–I10) said that there are only several tools that can meet their job needs. All participants stated that they generally do not consider programming when it comes to creating geographic articles in their daily lives, even if some of them have programming skills (I3–I6). In addition, all participants expressed that their understanding of professional software is quite limited because “*For me, one of the biggest challenges is learning how to use the software, and there are many configurations that need to be learned*” (I5–I7, I10). Most participants (I1–I3, I5–I9) emphasized that when writing geographic articles, they rarely use only one software or tool, as they typically start with a tool they are familiar with but may supplement it with another known tool if the first cannot fully complete the task of creating illustrations. In essence, they employ various methods and use all means at their disposal to accomplish the task of creating illustrations.

Regarding the general workflow for creating geographic visualizations, “*if there is data available, first import the data*” (I3, I6–I7, I9), then “*roughly determine the view range of the map and then proceed to draw the elements*” (I1–I10), and finally “*add other details manually, such as annotations*” (I2–I9). A few participants (I5, I10) described workflows that may omit one of the aforementioned steps or exhibit only minor variations but without changes in the overall sequence. Therefore, we posit that the workflow summarized above possesses a degree of universality.

Several participants (I3–I4, I9–I10) indicated that creating a geographic illustration often requires iterating through this workflow multiple times and involves tedious visualization layer integration. For example, they frequently need to use different software/approaches to separately create (1) map backgrounds, (2) visualizations generated from imported data, and (3) manually created visualizations. They then carefully align and merge these visualization layers using Adobe Illustrator or Photoshop while avoiding errors or misalignment in the process. They therefore expressed expectations for a coherent workflow that would allow them to determine the map, zoom to the appropriate area, draw points, lines, and areas, and add necessary annotations all in one go.

Furthermore, we have distilled the following themes from the interviews that include their pain points in the current workflows and their expectations, and summarized the design requirements for the desired tools.

1) *Pain Points and the Expectations: Frequent information comparison and application switching.* All participants stated that creating illustrations for text in geographic articles is a hassle because they need to “*match the text content, data, and maps*” (I1–I2, I6–I7, I9), which involves switching the focus back and forth between different applications, especially when they need to create and integrate multiple visualization layers using different software. Importantly, all participants agreed that maintaining high-quality authoring and minimizing rework hinges on improving the mental synchronization of text and illustrations. However, the current need to juggle multiple applications makes such synchronization difficult to achieve. Most participants (I1–I4, I7–I9) noted that after switching from one software to another and working in the new environment for a period of time, they often find themselves mentally

¹The formative interview was approved by the State Key Lab of CAD&CG, Zhejiang University.

drifting away from the content they were working on in the previous software. Therefore, regarding the desired tool, participants expressed an idea of linking articles and the illustrations to be created, to avoid the cognitive burden caused by repeatedly switching between them.

The workload and difficulty of manually creating geographic illustrations from text. Manual geographic illustration creation is difficult and troublesome, and the flexibility and diversity of textual fragment types pose challenges for text-to-geographic visualization. Through the participants' narratives of their own workflows, we uncovered that, even when texts contain identical or similar geographic information, variations in the specific textual descriptions may lead to differences in the content of the visualizations. For example, I1, who wrote news reports during an internship, as well as I7 and I8, all noted, *“even when writing articles about the U.S. elections, which often involve mapping the same states, the coloring method for these states differs based on the article’s intent and phrasing.”* I3 and I4 posited that even when referencing the same place names, the type and intended meaning of the textual fragment significantly impact the presentation style of geographic visualizations. To meet diverse narrative purposes when writing articles, it is necessary to master a wide array of visualization techniques, such as area selection, trajectory plotting, and color adjustment. Overall, manually creating geographic visualizations for articles is a tedious and challenging task. Participants thus expressed a desire for a tool capable of generating customized geographic visualizations based on textual content and tabular data. They believed such a tool would align with their common work requirements and potentially reduce their workload.

The need for generating text from geographic information. Although participants found creating illustrations more troublesome than writing, most of them expressed the need for transformation from geographic information to text. For example, I3 and I5 mentioned that when writing hiking plans, they need to *“brainstorm hiking routes while looking at maps,”* but *“the process of capturing information from maps and converting it to text is not easy.”* As another example, participants pointed out that news writing sometimes starts from data, and in such cases they often rely on AI for data summarization to inspire writing ideas. In these cases, participants typically feed geographic information or data to LLMs for text generation, yet prompt engineering is also a troublesome task, adding further disruption and interruption to already cumbersome workflows. Therefore, they hoped to find a tool that could help them write more easily, or provide prompts and inspiration for their article writing to address writing pain points, rather than only creating illustrations for the text.

2) *Design Requirements:* The interview results validated that our conceptualized generative approach could indeed enhance users' article writing process. First, many users lack a convenient and easy-to-learn tool for creating geographic illustrations, and a text-to-visualization generation method would typically be more intuitive and accessible. Second, the text-based generation approach naturally ensures that the produced illustrations are relevant to and aligned with the

textual content. Building upon this generative approach and incorporating the feedback from our interviews, we summarized the following design requirements for the expected tool:

R1: Generating geographic visualizations from the text. Users should be able to obtain corresponding geographic visualizations in a timely manner when composing textual content. To this end, the tool should accurately extract information from the text and generate aligned visualizations.

R2: Refining the generated visualizations. When the automatically generated visualizations differ from users' expectations or intentions, users should be able to adjust and modify the visualization results through convenient interactions, ultimately making them meet their expectations.

R3: Generating text from geographic information. Users should be able to leverage geographic information and visualization content from maps, or related geographic data, to easily convert them into textual descriptions, thereby helping to expand and inspire writing ideas.

R4: Interaction between text and visualizations. The tool should provide effective linking between text and corresponding visualizations, enabling users to easily identify and capture key information correspondences between them, thereby creating content more conveniently and effectively.

IV. PRELIMINARY STUDY

Prior to developing GeoAuthor, we conducted a preliminary study investigating 80 geographic articles from various sources, including The New York Times, BBC, National Geographic, and RedNote [76], to analyze and summarize the relationships between geographic articles and the visualization forms of corresponding illustrations; the full list of articles is provided in the supplementary materials. Our main findings are summarized as follows.

Geographic Visualization Breakdown. We decompose the final presentation of illustrations collected from the articles into two primary layers: map background and graphics drawn on the map. The **map background** involves common map types (standard, satellite, traffic, etc.) zoomed to different scales and areas based on the context. Textual intent of the article influences the selection of the map background. 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 [77]: **point**, **line**, and **area**. The second category comprises **annotations** derived from these primary elements, such as text labels, bounding boxes, and similar elements. 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 consider these to be the most common visualizations in geographic articles.

Furthermore, textual content influences the display form of geographic elements, specifically affecting both the determination of component types and their subsequent visual rendering. Regarding type determination, a place name, for instance, could reasonably be displayed as either a point location or an

area on a map when considered in isolation. However, when the purpose of the contextual narrative is taken into account, we can infer which type is more appropriate. As for the specific rendering, for instance, once it has been established that a piece of text should generate a trajectory, the decision to represent it as a straight line or a curve depends on the specific textual description.

Typical Visual Encoding Methods. In geographic article illustrations, visual encoding methods are typically straightforward and can be inferred to a considerable degree from the textual content. Considering this, we can broadly categorize the visual encoding in these illustrations into two main types:

Categorical encoding involves assigning different visual appearances based on categories. Objects of the same type are encoded consistently, while objects of different types are encoded differently. For example, if the text mentions that certain areas have one characteristic while other areas have another, different colors can be assigned to different regions on the map based on these categories.

Numerical encoding involves encoding values for different objects as a continuously changing function rather than the discrete approach of the first case. For example, a political news article mentions election support rates across communities. On the map, the redder the communities, the higher the Republican support rate, while the bluer they are, the higher the Democratic support rate.

For more details about the preliminary study, such as the research process, the detailed descriptions and examples for the visualization types, please refer to Appendix B.

V. GEOAUTHOR

We developed our system GeoAuthor based on the findings from formative interviews and the preliminary study.

Implementation. GeoAuthor is built on a client-server architecture. The frontend is implemented using Vue.js and the backend using Python Flask. The LLM runs on a server equipped with 192-core AMD EPYC 7763 processors, 384 GB RAM, and an NVIDIA A100 (80GB) GPU.

A. Geographic Visualization Generation

GeoAuthor provides features to facilitate immersive and easy creation of geographic illustrations for authors (R1).

While writing the article in a text box (Figure 2B), users can split the article in a text box into two text boxes by clicking or merging the text in different text boxes by dragging. The article from a text box can be leveraged to generate the geographic visualization (Figure 2C). Specifically, GeoAuthor sends the article to the LLM², which returns formalized content in JSON format. By parsing the JSON content, GeoAuthor generates the geographic visualizations (R1). The prompts for generating the JSON content from the text are carefully compiled using Few-Shot Learning [78] and contain five key components: instructions, visualization type definitions, several examples, analysis approach and precautions, and the output format. The JSON grammar includes map background and zoom

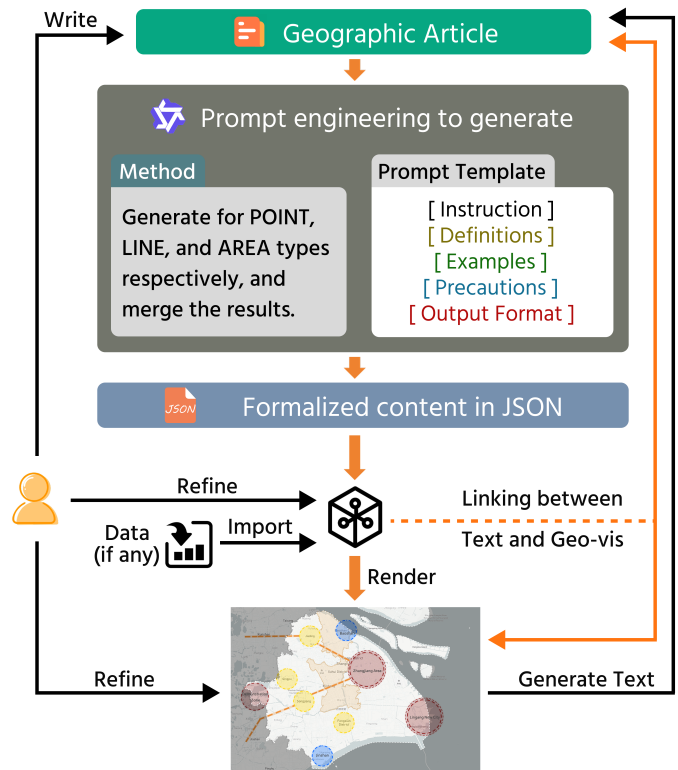


Fig. 1. The overall workflow of GeoAuthor. Users write articles, which generate JSON grammar through LLM prompt engineering, and visualizations can be rendered by parsing it. Then, users can refine the generated visualizations to better align with user intent. GeoAuthor can also utilize geographic information to generate text to assist writing. Overall, GeoAuthor links geographic text and geographic visualizations, effectively facilitating article authors in creating illustrated geographic articles.

scale, eight visualization types (identified in the preliminary study), where each type contains the index of each element in that type, the involved place name sets or sequences, visual encoding, source text location, and other attributes. The overall workflow of GeoAuthor is presented in Figure 1, and a detailed prompt template and a structural example of the JSON grammar can be found in Appendix C.

In addition, users can specify whether to use categorical or numerical encoding types. For categorical types, GeoAuthor automatically categorizes geographic objects based on article semantics and assigns different visual encodings to objects in different categories, while applying uniform visual encodings to objects within the same category (e.g., Figure 2C2). For numerical types, users only need to assign visual encodings (such as color and height) to geographic objects representing the two extreme values, and GeoAuthor automatically applies transitional colors to other geographic objects based on their corresponding data values (e.g., Figure 2C1).

B. Auxiliary View

The auxiliary view (Figure 2A) serves as a bridge between text and visualization, explaining the visualization approaches as understood by the LLM. Considering that the visualization generated from the article in a text box can potentially be decomposed into multiple elements, the auxiliary view

²The LLM we used is Qwen2.5-7B-Instruct.

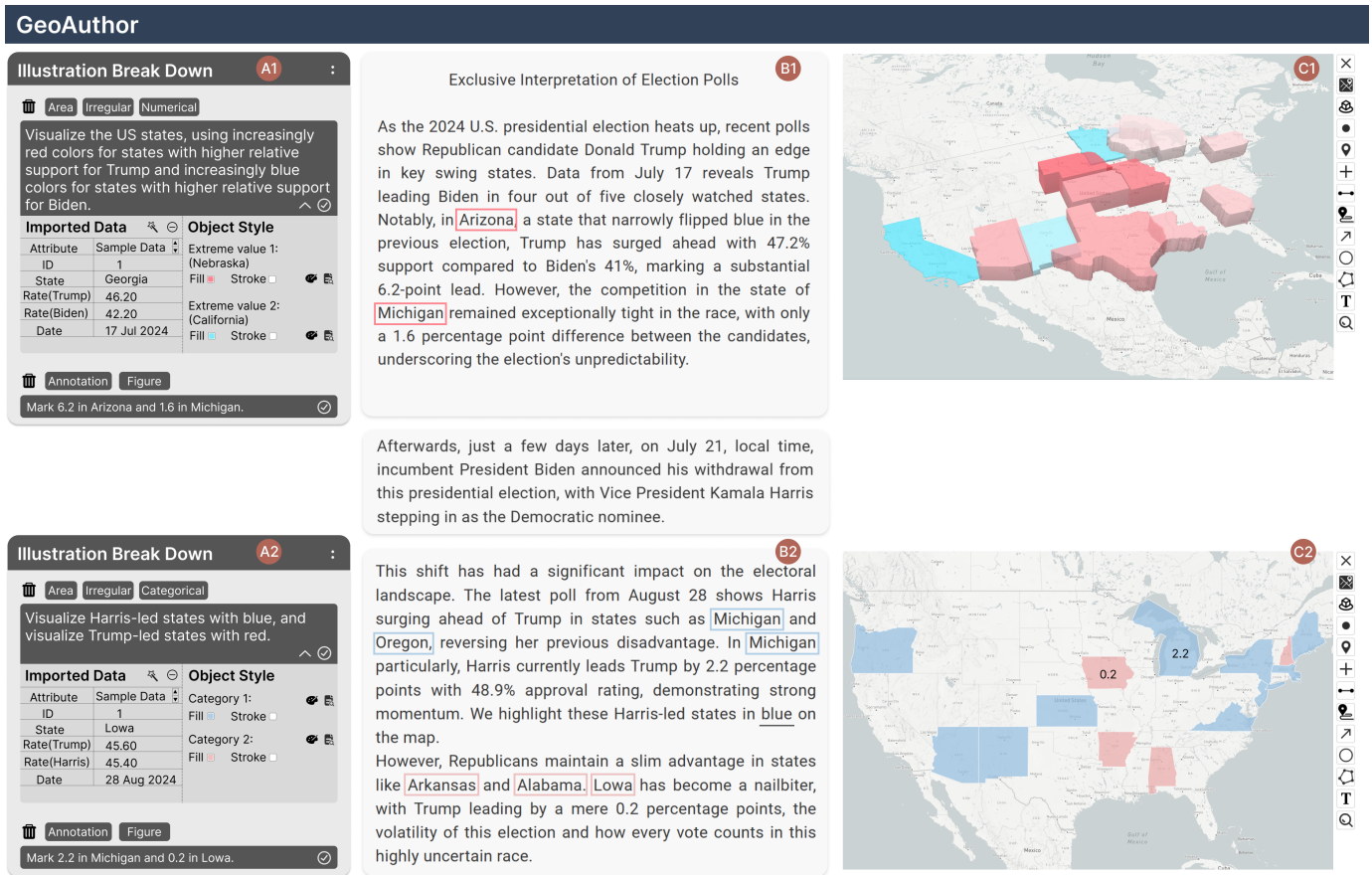


Fig. 2. Interface of GeoAuthor: (A) The auxiliary view communicates the article and the generated visualization; (A1) The labels “Area,” “Irregular,” and “Numerical” indicate that the visualization type is the irregular area and the numerical encoding is applied, because each state is an irregular region, and its color needs to be encoded based on the support rate values; (A2) The labels “Area,” “Irregular,” and “Categorical” indicate that the visualization type is the irregular area and the categorical encoding is applied, because Harris-led states need to be drawn in blue and Trump-led states need to be drawn in red; (B) The text box that contains an article snippet; (C) The generated geographic visualizations. This figure also demonstrates our usage scenario I.

represents each element through four components: labels, explanatory text, imported data, and object styles. The labels provide a summary of the visualization type and its visual encoding. Beneath the labels, the explanatory text offers a detailed explanation of the visualization type and encoding rules derived from the article and the imported data. Further below, the imported data and object styles are presented. The initial styles are automatically recommended based on article semantics, textual sentiment, and common sense.

GeoAuthor supports data import in CSV format. The imported data can either be manually defined by the user to directly generate visualizations or combined with the article to create visualizations:

- For the former, users can select and specify meanings for data columns, with supported meaning types including place names, latitude/longitude coordinates, GeoJSON, and numerical columns for visual encoding (optional 1–2 columns). GeoAuthor can thus understand the data and utilize it to generate visualizations. For example, in Figure 2C1, for this type of numerical visualization, articles often do not explicitly list all geographic names and their associated data, making it necessary to leverage the imported data to generate the visualization.

- For the latter, it accounts for cases where the visualizations in an article’s illustrations may not fully correspond to the textual content. For instance, in Figure 2A2, B2, and C2, the article categorizes states into two types but does not exhaustively list all state names within each category. By sending both the article and data to the LLM and requesting categorical visual encoding, the LLM can automatically complete the classification based on article semantics and specific data. In this example, states are divided into two categories according to whether Democrats or Republicans have higher approval ratings, and are rendered with different colors respectively.

C. Geographic Visualization Refinement

Although GeoAuthor generally performs well, there may be instances where the generated visualizations are inadequate or deviate from the user’s intent. GeoAuthor therefore provides the following approaches for visualization refinement (R2).

- 1) *Modifying the Auxiliary View:* All content in the auxiliary view can be freely modified according to the user’s intent. Users can directly modify labels in the auxiliary view to change visualization types, edit explanatory text to alter visualization objects, presentation formats and styles, and

adjust object style panels to directly modify visual encodings. GeoAuthor automatically updates the geographic visualization based on these changes (R2). For categorical types, when modifying the style of a geographic object, objects in the same category will change together in a bound manner. For numerical types, when modifying the visual encoding of objects with extreme values, the visual encodings of other objects will change accordingly.

2) *Directly Using the Map*: By clicking on any element of the generated visualizations on the map, a popup window will appear, allowing users to manually modify its style, such as point size, line thickness, and color (R2). Users can also utilize the toolbar on the right side of the map to add annotations to the map and switch between different map backgrounds, including standard maps, satellite maps, traffic maps, and dark mode maps. The “+” button in the toolbar allows users to import an icon and add it to the map. The imported icon can also be used for the “other icons” type visualization, one of the eight visualization types mentioned previously.

3) *Disambiguation*: GeoAuthor features disambiguation of locations, regions, and trajectories (R2).

Location/Region Disambiguation. Occasionally, a place name yields multiple search results on the map, such as universities with multiple campuses or identically named roads and buildings, requiring toponym disambiguation. If a displayed location/region does not match the users’ intended reference, they can right-click on it on the map, and GeoAuthor will automatically switch to an alternative location/region with the same name from the search results. This process repeats iteratively until the desired location/region is identified. Since GeoAuthor automatically extracts search scope based on text semantics, such as limiting “Statue of Liberty” searches to the U.S., the search results typically do not exceed a few options.

Trajectory Disambiguation. Similarly, users can disambiguate trajectories composed of a series of key locations interconnected in various ways. The specific connection between several locations, whether using straight lines or a meaningful curve, is determined by the textual description, which may include references to modes of object movement or human transportation. However, when text lacks clarity about these aspects, users can click on the trajectory to access disambiguation options, allowing them to refine the visualization of the corresponding trajectory.

D. Text Visualization

If a geographic description in the article (e.g., a place name) is color-coded in the geographic visualization, a rectangular color box will highlight the corresponding geographic description in the article (R4), as shown in Figure 2B and Figure 6. The color of the rectangular box in the article is tightly bound to the color of the corresponding geographic element on the map; any change in the latter will automatically trigger a corresponding change in the former (R4). Conversely, the highlighted text can be clicked to open a popup label that displays its visual encoding on the map and allows direct modification (R4). This avoids users having to repeatedly cross-check and compare text and geographic visualizations,

allowing users to focus on writing and to reduce mental effort. This is also another approach for visualization refinement: adjusting visualization styles in the map directly from text.

Similarly, keywords in the article related to visual encoding are also closely bound to the geographic visualization (R4). For example, if the article mentions that a certain area is rendered in blue on the map (Figure 2B2), GeoAuthor will automatically color that area blue when generating the visualization and underline the word “blue” in the article. If the user later modifies the color of that area on the map, the word “blue” in the article will automatically change accordingly.

E. Text Generation

GeoAuthor can generate text from both geographic visualizations and geographic data (R3).

1) *Text Generation from Geographic Visualizations*: First, clicking any element in the generated visualizations or annotations on the map opens a popup where users can generate multiple descriptive text options for the selected element (R3). These texts can be dragged and dropped into the article.

Second, when the users right-click anywhere on the map background, a label containing the place name of that point will appear. By dragging this label into the text box, GeoAuthor will automatically continue writing based on the existing text and the place name in the label (R3).

Third, GeoAuthor is capable of generating spatial relationship descriptions. Users hold Shift and click visualization elements or map locations, and the system generates textual descriptions of the selected elements’ relationships (R3). For example, “Beijing and Shanghai are 1,000 kilometers apart... They are respectively the political and financial centers of China...”, or “the highway facilitates travel between the surrounding cities A, B, and C.”

The implementation of the above three features involves sending article context and the attributes of the selected elements to an LLM, ensuring generated text maintains coherent style and content with the current article. Moreover, in the second and third features above, when users select arbitrary locations on the map rather than existing elements, the system converts these locations into textual place names by inferring the appropriate administrative region level based on the current zoom level before transmitting to the LLM. For instance, if a user selects a location with coordinates corresponding to “New York University,” the system determines whether the user’s intent refers to “New York University” as a POI (Point of Interest) or a broader administrative region such as “New York City,” depending on the current map zoom level. The zoom levels are categorized into four types: national level, state level, city level, and POI level.

2) *Text Generation from Geographic Data*: GeoAuthor can analyze patterns in the users’ imported data and, in conjunction with existing text, continue writing an analytical summary of the data (R3). The implementation method is also based on LLM, sending the visualization screenshot, data, and article together to the LLM to generate textual descriptions.

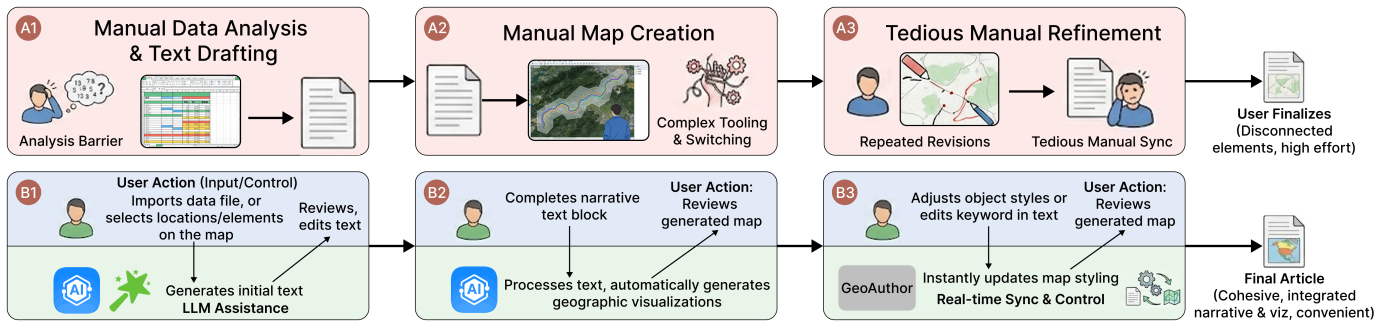


Fig. 3. Workflow comparison for authoring geographic articles. (A) Traditional workflow: Without LLM assistance, it requires (A1) manual data analysis/text drafting, (A2) manual map creation, and (A3) tedious manual refinement steps, resulting in high effort. (B) GeoAuthor’s approach: With LLM assistance, (B1) users can seek support from the LLM when writing text, then review and revise; (B2) Users only need to immerse themselves in writing, while geographic visualizations can be generated with one click, eliminating the cumbersome and complex tool usage and switching; (B3) Users finally make slight adjustments to the visualizations to achieve satisfactory results. The entire process is more convenient and efficient.

F. Consistency and Visual Links

Consistency. The intermediate grammar in JSON format generated in the visualization generation process establishes a bidirectional linkage between the text and the geographic visualizations (R4). Based on this, regardless of whether users make modifications to the article, auxiliary view, or map, GeoAuthor automatically updates other parts that need to change, consistently maintaining coherence among all three components. The auxiliary view itself is also a manifestation of the link between text and visualization, allowing users to more conveniently revise the generated visualizations.

Visual Linkage. There are effective linkages between text and geographic visualizations. When a user hovers the mouse over any highlighted text in the article, the corresponding geographic element on the map will be highlighted while other elements on the map are dimmed; Conversely, when the mouse hovers over a geographic element on the map, a label displaying the element name appears, and the corresponding article segments in the text box are also highlighted (R4). In this way, users can easily identify the correspondence between keywords in the text and geographic elements on the map.

VI. USAGE SCENARIOS


This section presents two usage scenarios, demonstrating GeoAuthor’s capabilities in enabling users to create geographic articles conveniently and effectively. Figure 3 shows a comparison between the GeoAuthor approach and the traditional workflow. We also present additional examples of GeoAuthor’s creation results in Appendix A.

A. Usage Scenario I

Emily was a journalist working at a news agency, using GeoAuthor to write her news article about the U.S. election.

She first imported into GeoAuthor a dataset of state support rates from U.S. election polls (Figure 2A1) that was provided by colleagues from the data department. Based on this, she completed the first paragraph of her news article (Figure 2B1), which discussed the polling situations for Trump and Biden across various states. She then had GeoAuthor generate a numerical type geographic visualization for this passage, with

states having relatively higher Republican support rates colored more red, and states with relatively higher Democratic support rates colored more blue. Not satisfied with the initial colors, she modified the colors of the two states with extreme color values in the auxiliary view (i.e., the states with the highest relative support rates for Trump and Biden respectively), and GeoAuthor immediately updated the colors of all states on the map based on this change (Figure 2C1).

Afterward, she wanted to analyze the changes in polling dynamics after the incumbent president withdrew from the race, such as which states still maintained their advantage and which states experienced reversals. However, faced with the boring data table, she did not know where to start with analysis and writing. So she first wrote the opening sentence of a new paragraph, then clicked the  button in the “Imported Data” panel to have the system generate subsequent text based on the imported data. Finding parts of the generated text satisfactory, she kept the general content and made some minor modifications, ultimately completing the entire paragraph (Figure 2B2). Since this required categorical visualization, she had GeoAuthor automatically divide the states discussed in the article into two groups based on the narrative intent and applied different colors on the map (Figure 2C2). Eventually, she completed the composition of the entire article and felt satisfied with both the text and the illustrations.

B. Usage Scenario II

John, a college student who enjoyed outdoor activities and socializing, often posted on campus forums to find people to go hiking together. Now he had opened GeoAuthor.

John began to write his weekend outing plan. As he wrote, he checked the map, brainstorming for new and interesting places to go. He noticed a green park, so he clicked on that location, and a label containing the place’s name, “Imperial Palace”, appeared on the map (Figure 4). John then dragged this label into the text box and GeoAuthor immediately continued writing based on the existing text and the place name he had just dragged in. John was satisfied with the system-generated content and then wrote more text to complete the entire plan and asked GeoAuthor to generate the visualization.



Fig. 4. John wrote an opening for his outing plan. He searched for travel destinations on the map, and noticed an interested region “Imperial Palace.”

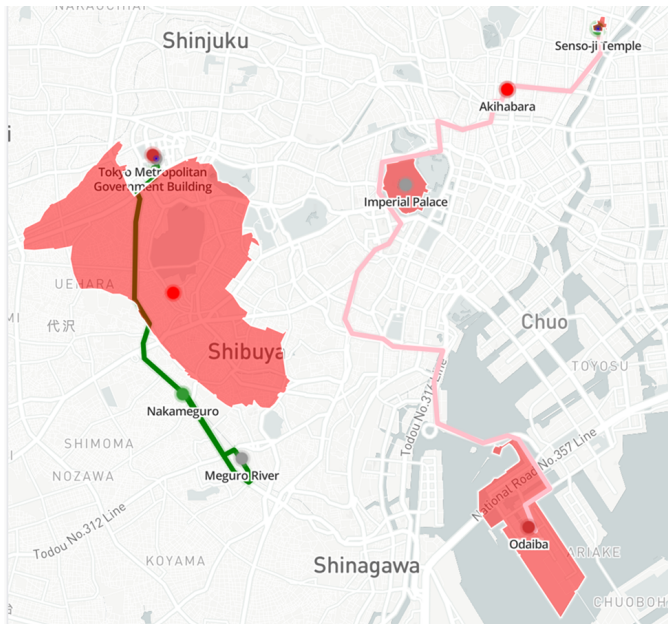


Fig. 5. After John completed the entire outing plan, in GeoAuthor’s initial visualization generation, although it included trajectories for the two-day itinerary, it contained other redundant elements, making the entire map somewhat cluttered.

Although GeoAuthor intelligently generated two trajectories to represent the two days’ travel routes described in the text, it had visualized all the place names mentioned in the text, making the map appear cluttered and failing to highlight key points (Figure 5). Consequently, he used the auxiliary view to delete some elements that did not need to be drawn and set the dot type elements as marker type. After saving these changes, GeoAuthor regenerated the visualization.

This time, John was satisfied with the generated result. He then modified the styles of geographic elements on the map according to his personal preferences and added a few annotations (e.g., speech bubbles labeled “Day 1”, “Day 2”), thereby completing the creation of the entire article (Figure 6).

VII. EVALUATION

In this section, we validate the usability of GeoAuthor through system evaluation and a user study.

A. System Evaluation

1) *Generating Visualizations from Text*: We quantitatively evaluated the accuracy of GeoAuthor in generating visualizations and the response time.

Benchmark. To evaluate the accuracy and runtime of GeoAuthor’s automated text-to-visualization generation, we constructed a benchmark of 50 evaluation cases, which is provided in the supplementary materials. Each case consists of a piece of geographic text and its corresponding ground-truth visualization specification in JSON format, which is used to assess whether GeoAuthor correctly parses the text into the expected JSON output.

The benchmark construction adopts a semi-automated approach, using an interface we developed to build geographic texts of different themes containing different visualization types and their corresponding correct JSON parsing results. In the future, this interface can facilitate convenient expansion of the scale of the test benchmark. For detailed processes and interface screenshots, please refer to Appendix D.

Accuracy of Generated Visualizations. When evaluating test results, we considered that GeoAuthor’s output might sometimes contain more optional elements than the ground truth in the benchmark. For example, in a travel blog, the most important aspect might be drawing the trajectory, while rendering certain regions could be optional—this involves subjectivity. Therefore, we judged such cases as correct outputs.

Test results show that 44 (88%) cases produced correct outputs, while among the 6 incorrect cases, 5 had the error type “missing elements” and 1 had the error type “elemental chaos.” This indicates that the system can generate fundamentally correct visualizations in most cases. These rare wrong cases typically occur when text descriptions are very vague or chaotic, making it difficult for the LLM to clearly infer the text author’s visualization creation intent. For the few incorrect cases, users can revise them through interactive methods in GeoAuthor to ultimately achieve a satisfactory outcome. Detailed test results for accuracy can be found in the supplementary materials.

Response Time. When testing on the aforementioned benchmark, we also measured GeoAuthor’s runtime performance. Text-to-visualization generation constitutes the primary latency component in GeoAuthor, encompassing JSON content generation followed by visualization rendering. We calculated the running time across the 50 cases in the benchmark, with an average result of 23.83 seconds, where the LLM-powered text-to-JSON stage took 8.46 seconds, and rendering the JSON results on the map took 15.37 seconds. The map rendering stage exhibits extended execution time due to our current implementation approach for GeoAuthor: For each geographic entity in the JSON results requiring visualization (e.g., “Louvre”), GeoAuthor performs real-time OpenStreetMap API queries to retrieve corresponding latitude-longitude coordinates or GeoJSON data for subsequent rendering on the map. The combination of API rate limitations (maximum one request per second) and network latency results in extended runtime for this stage. Future optimizations, including pre-cached GeoJSON libraries for frequently referenced locations, could significantly improve computational efficiency. Detailed runtime results can be found in the supplementary materials.

2) *Quality of Generating Text*: To evaluate the quality of text generated by GeoAuthor, we ran GeoAuthor 50 times across the four text generation features described in Section

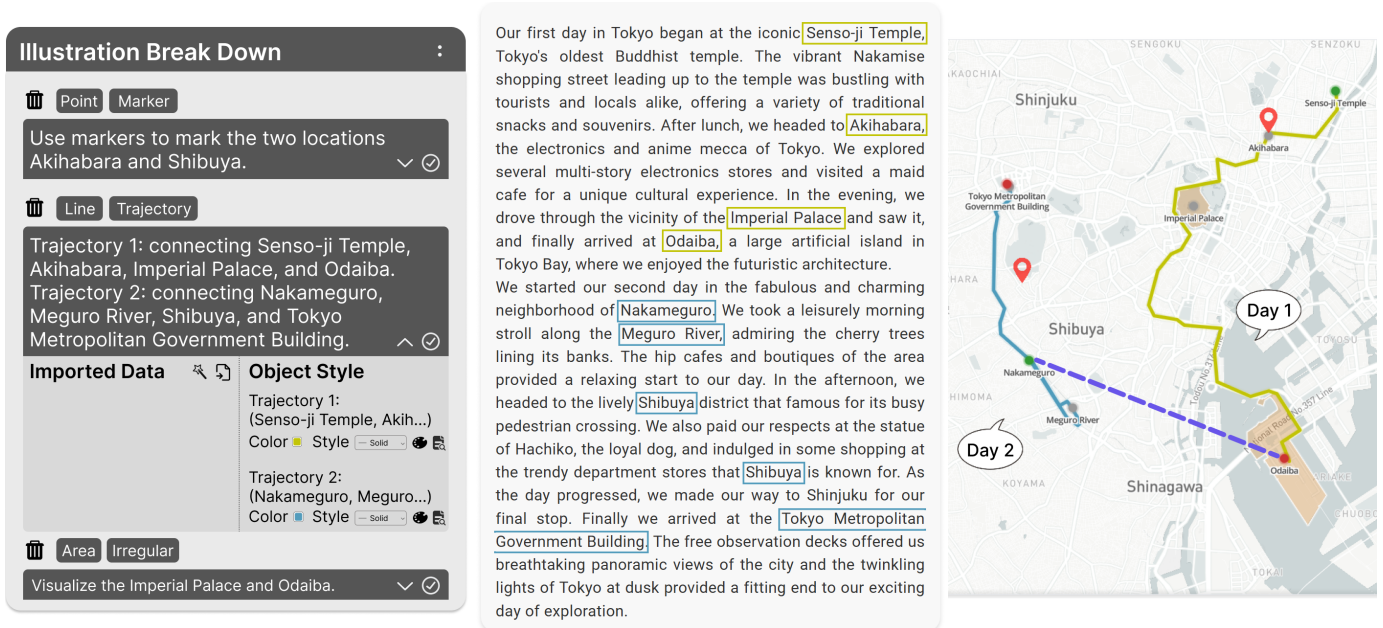


Fig. 6. The interface state when the geographic article and illustration creation is completed in usage scenario II. The auxiliary view on the left provides detailed explanations of the markers, trajectories, and irregular regions rendered in the map. The map on the right displays the geographic visualization creation results. The highlighted text colors in the central text box remain consistent with the corresponding trajectory colors in the map. It should be noted that entities such as “Odaiba” are visualized not only as part of the “trajectory” type in the map but also rendered as “irregular region” type visualizations. Their text highlighting colors can correspond to either the former or the latter’s color in the map, which can be manually selected by the user.

V.E. The generated outputs were manually assessed by two authors (the first and second authors) independently, and any disagreements were resolved through cross-checking and discussion. The evaluation considered three criteria: accuracy (whether the generated content is factually correct and free from geographic errors or hallucinations), fluency (whether the wording and sentence-level expression are natural and coherent), and logic (whether the content is reasonable and well organized, with clear progression and appropriate connections between ideas).

Among the 50 outputs, 35 (70%) showed no issues on any of the three criteria. 10 outputs (20%) exhibited accuracy issues, mainly hallucinated or geographically incorrect content. 4 outputs (8%) showed logic issues, and one output (2%) exhibited both accuracy issues and logic issues. Overall, the generated text quality is satisfactory and supports GeoAuthor’s goal of helping users expand their writing ideas. Detailed results are provided in the supplementary materials.

B. User Study

To validate the usability of GeoAuthor, we conducted a comparative user study³ involving 12 participants (P1–12), all of whom have experience in geographic article creation. None of them had been involved in the development of GeoAuthor.

1) *Baseline*: We developed an ablated version of GeoAuthor as our baseline. This baseline removes GeoAuthor’s core features, including text-to-visualization generation, text generation, and the linking between text and geographic visualization. In the baseline, users can create geographic illustrations

through the toolbar on the right side of the map while writing articles; The auxiliary view only retains data import, and the elements added to the map and their styles. This baseline ensures that all tasks in the user study can be completed while simulating typical workflows. Through this ablation study, we aim to on the one hand demonstrate GeoAuthor’s usability, and on the other hand validate the effectiveness of the linking between text and visualization.

Justification. We did not include existing mainstream professional tools as baselines because, to the best of our knowledge, no mature tool is directly comparable to GeoAuthor. In existing practice, users typically need to combine multiple tools, resulting in a longer and more fragmented authoring workflow. Moreover, these tools generally lack automated generation features and instead require manual creation of both visualizations and text. Therefore, a direct comparison would be difficult to make fair, given the substantial differences in workflow and level of automation. We thus used an ablated version of GeoAuthor as the baseline to more directly evaluate the contribution of its core features.

2) *Procedure*: Each participant went through the following procedure:

Introduction. The facilitator introduced the background and demonstrated the features of both GeoAuthor and the baseline through several use cases. Participants were then given warm-up time to gain hands-on experience with both GeoAuthor and the baseline.

Semi-Structured Geographic Article Completion Task. Participants were asked to complete two semi-structured geographic writing tasks sourced from real online articles—one from a New York Times news article and another from Red-

³The user study was approved by the State Key Lab of CAD&CG, Zhejiang University.

Note. Portions of these articles were pre-edited and truncated to resemble semi-finished drafts. The facilitator provided writing and illustration instructions, and participants completed the tasks using both GeoAuthor and the baseline. The order of tool usage was counterbalanced to ensure half of the participants used GeoAuthor first while the other half used the baseline first. Throughout the entire process, both tools were anonymized to prevent bias-inducing guidance. For example, participants were not informed which tool was the baseline, and the facilitator avoided any biased language that might influence their judgment.

Questionnaire and Interview. Participants completed a System Usability Scale (SUS) [79] questionnaire and a custom system effectiveness questionnaire for both GeoAuthor and the baseline, using a 5-point Likert scale. The facilitator then conducted semi-structured interviews to gather insights on the utility and limitations of system features.

3) *Quantitative Results: The SUS Questionnaire Results.* All participants successfully completed the tasks. The detailed results of the SUS questionnaire for GeoAuthor and the baseline are presented in Figure 7. Participants rated GeoAuthor and the baseline with SUS scores of 86.67 and 72.5, respectively, with the former exceeding the 80.3 threshold for the top 10% and the latter surpassing the 68-point threshold for the top 50% [79]. Based on SUS factor research [80], we calculated GeoAuthor’s usability and learnability scores as 86.2 (Q1–3, Q5–9) and 88.54 (Q4, Q10), respectively, and also calculated the baseline’s usability and learnability scores as 69.27 and 85.42, respectively. These results indicate that while the baseline demonstrates high learnability but only moderate usability, GeoAuthor’s text-visualization linking capabilities enable excellent usability that significantly surpasses the baseline. This therefore validates the effectiveness of linking text and visualization for geographic article creation. In the interview phase, participants indicated that GeoAuthor, due to its “*integration of rich and effective functionality*” (P3), “*requires learning and exploration*” (P12), but this “*is just getting familiar with some interaction operations, does not require learning much new knowledge, so there is no significant learning burden*” (P12). P10 stated that even though the baseline “*is close to my usual work approach*”, he still found the baseline “*awkward*” when compared to GeoAuthor. Additionally, while the vast majority of participants gave positive evaluations to GeoAuthor, P2 gave only 1 point in Q2, stating “*I cannot remember so many functions at once.*”

The System Effectiveness Questionnaire Results. The results of the custom system effectiveness questionnaire are presented in Figure 8A. We assessed the internal consistency reliability of the questionnaire using Cronbach’s α . The results showed good internal consistency for both GeoAuthor ($\alpha = 0.812$) and the baseline system ($\alpha = 0.851$). Moreover, GeoAuthor significantly outperformed the baseline across seven questions (Q1–4, Q6, Q9–10), with only three questions (Q5, Q7, Q8) showing comparable results. Detailed results in the Student’s t-test are provided in Appendix E.2. From the Q5 results, it can be seen that GeoAuthor and the baseline have comparable user-friendliness. Regarding the Q7 and Q8 results, P1 explained that “*both tools produce accurate and*

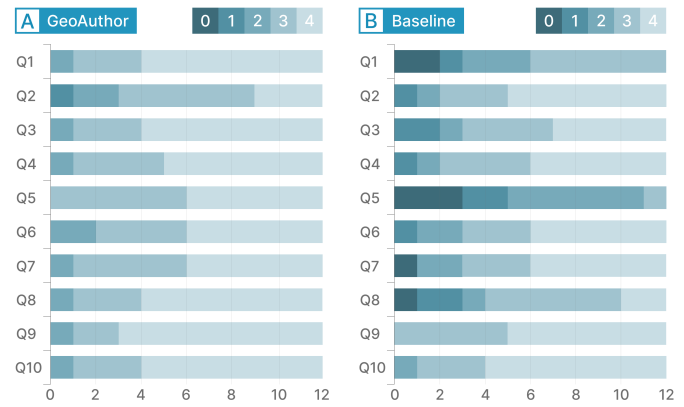


Fig. 7. Detailed results of the SUS questionnaire for (A) GeoAuthor and (B) the baseline.

aesthetically pleasing illustrations, with no particularly poor outcomes.” In the interview phase, participants indicated that the greatest difference between GeoAuthor and the baseline is that GeoAuthor is “*more convenient to use*” (P1, P3, P6) and “*more efficient*” (P3), while the baseline’s functionality “*can be replaced by other apps*” (P3). Additionally, P3 mentioned that in the past, she would “*use mapping software to manually plot points and manually draw trajectories,*” while the paths drawn by GeoAuthor are “*more reasonable than the gold standard from RedNote app.*”

The Completion Time. The completion time for each task (including system response time) is shown in Figure 8B. We used Student’s t-test for analysis: In Task 1, GeoAuthor’s completion time was significantly shorter than the baseline ($t(11) = -6.0914$, $p < 0.001$, Mean Difference MD = -3.3292, 95% CI = [-4.5321, -2.1262], Cohen’s $d = -1.7584$); In Task 2, GeoAuthor’s completion time was slightly shorter than the baseline but not statistically significant ($t(11) = -0.5748$, $p = 0.5770$, MD = -0.6683, 95% CI = [-3.2273, 1.8907], Cohen’s $d = -0.1659$). We believe this is because Task 1 was relatively simpler, while Task 2 involved more locations with complex sequential relationships, making the LLM more prone to confusion and resulting in users needing more time to revise. Additionally, there was an outlier in Task 2 (Figure 8B), which we determined was caused by network latency after reviewing user study recordings.

4) *Qualitative Results:* We summarize the interview findings into the following conclusions, including GeoAuthor’s effectiveness and limitations:

GeoAuthor can effectively assist in creating geographic visualizations (R1, R2). Several participants (P1–2, P4–6, P8, P10) indicated that they considered GeoAuthor’s most practical feature to be text-to-visualization generation, especially when “*the article is complex and involves many place names and locations.*” Without GeoAuthor, “*geographic illustrations can still be created, but efficiency would be much lower*” (P3, P5). Participants were generally satisfied with the accuracy and quality of generated visualizations, believing they eliminated the tedious process of geographic illustration creation. Additionally, participants indicated that the categorical and numerical visual encoding provided by GeoAuthor

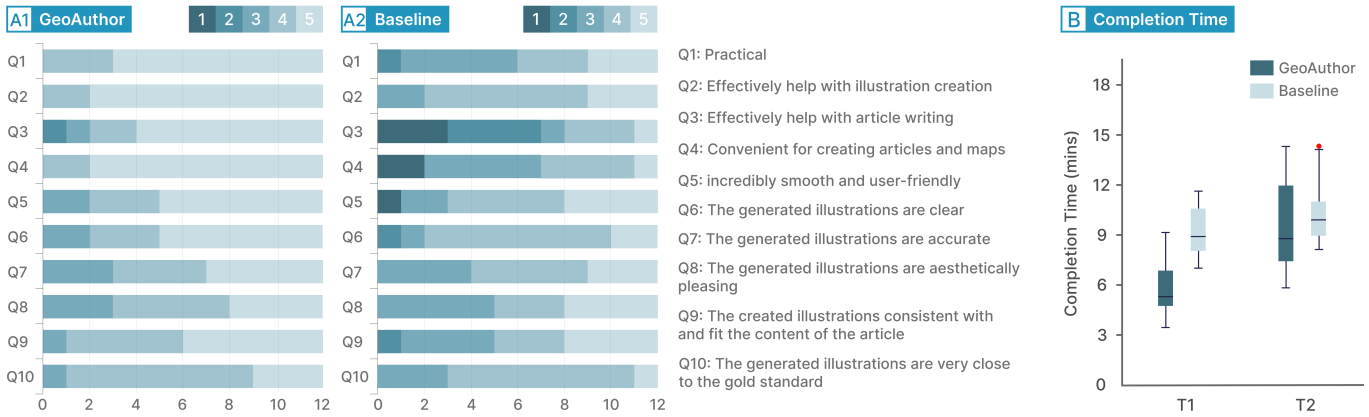


Fig. 8. (A) The results of the custom system effectiveness questionnaire for GeoAuthor and the baseline, where 1 = strongly disagree and 5 = strongly agree; For Q10, we provided participants with the illustrations from the original sources of the articles used in the tasks as the gold standard during the questionnaire phase. (B) The completion time (minutes) for GeoAuthor and the baseline in two tasks.

is very practical, especially P3 who mentioned she “*had previously manually modified the styles of over 100 points.*” Even though GeoAuthor’s generated visualizations sometimes “*derive from machine intent and deviate from human intent*” (P2, P6), participants unanimously indicated that they could still achieve satisfactory results through visualization refinement features such as auxiliary views. Additionally, while GeoAuthor showed measurable response latencies when generating visualizations from articles, participants consistently affirmed that it offered greater efficiency and convenience compared to the baseline, with response times deemed acceptable. For instance, P5 considered GeoAuthor’s response time to be “*controllable*” and noted “*dynamic effects of progressive visualization rendering during this waiting process.*”

However, P7 raised a concern that when place names in articles become complex and their sequence and descriptions are disorganized, the accuracy of generated visualizations may decline. While users can address this issue through refinement, it may impact efficiency. To further improve accuracy, we could consider using better LLMs or fine-tuning LLMs specifically for our text-to-geographic visualization tasks.

GeoAuthor’s text generation capability is exciting and valuable (R3). Several participants (P1, P4–7, P10) found GeoAuthor’s text generation capability practical, noting that “*the generated text is very appropriate*” (P7), “*it is friendly for people with poor writing skills*” (P6), “*it can generate a rough draft that you can then polish and modify*” (P5), and “*it saves workload, especially when summarizing and extracting insights*” (P10). Overall, participant feedback validated that GeoAuthor’s text generation capabilities can indeed provide inspiration or reference for writing, reducing tedious typing work, which may disrupt the overall workflow and writers’ creative thinking. Several participants (P2, P6) indicated that they often follow a workflow of creating visualizations first and then writing, and GeoAuthor’s ability to generate textual descriptions or summaries from geographic content was found “*interesting and exciting*” (P6).

However, P3 and P8 suggested extending GeoAuthor’s map-to-text generation, which is currently basic, to better support

data-first authoring workflows. In addition, a few participants noted that “*the generated text does not necessarily fully align with user intentions*” (P2) and “*the writing has a strong AI flavor and lacks personalization*” (P3). P8 and P12 therefore suggested adding an LLM-powered natural language interface to meet writers’ personalized writing and editing needs.

The linking between text and geographic visualization effectively reduces cognitive burden and facilitates authoring (R4). Participants expressed appreciation for the linking between text and visualization, as it “*reduces interactive burden*” (P6) and “*helps focus attention, especially when there is a lot of text*” (P12). During the authoring process, the interaction between text and visualization is “*more intuitive and convenient than traditional GIS software, especially when dealing with unfamiliar regions*” (P3). P6 indicated that this linking and correspondence can help with checking and ensuring nothing is missed, making users feel more confident, especially when dealing with a large number of place names. P11 stated that her most appreciated feature was directly modifying corresponding colors in the map through text, as this is “*intuitive and WYSIWYG (what you see is what you get),*” while reading auxiliary views involves “*reading machine-generated configurations,*” and they contain “*too much text that requires time to process.*” However, P10 felt that this feature was not the most effective, although he acknowledged it was still useful. He noted that when text is lengthy and some place names appear multiple times, he would consider directly modifying styles on the map rather than approaching it through the text.

This linking between text and geographic visualization should not only exist during the authoring process, but should also be preserved in the final work presented to readers, enhancing readability and reader engagement. For example, P4, an urban planner, believed that GeoAuthor’s application scenarios “*are not limited to writing, but can also extend to reading, such as when reading articles about urban renewal that contain many names of old residential communities,*” where GeoAuthor’s text-to-visualization capabilities and the linking between them can “*greatly improve work efficiency.*”

This also demonstrates the generalizability of our method.

VIII. DISCUSSION

In this section, we discuss the implications, limitations, and future work of our research.

Implications. GeoAuthor not only facilitates the integration of text and visualization in geographic article creation but also strengthens the interaction between narrative and spatial representation. By dynamically linking textual content with visual outputs, it allows for iterative refinement, where changes in one medium can directly inform and enhance the other. This bidirectional linkage addresses a critical gap in traditional workflows, where authors often struggle to ensure alignment and consistency between descriptive text and its corresponding visuals. Overall, we believe GeoAuthor demonstrates novelty and utility, effectively addressing the pain points and needs of article authors while promoting immersive creation.

Limitations. There are limitations discussed as follows. Firstly, GeoAuthor lacks support for overly complex visualizations. Its supported visualization types are based on classifications from our survey of geographic articles. Ignoring software learning curves and the effort or time needed to create visualizations, authors could theoretically use any tool to create more diverse visualizations. For example, geographic articles on news sites like The New York Times sometimes feature highly complex or refined illustrations, though these often demand significant time and expertise to produce. GeoAuthor therefore has its advantages, namely that it facilitates more convenient and efficient creation when dealing with common visualization types. Secondly, GeoAuthor falls short compared to professional GIS software in handling complex geographic data visualization. Designed to link text and visualization for article writing, it focuses on quickly generating visuals from textual descriptions but lacks advanced GIS features like complex spatial analysis, large dataset management, or custom map projections. Its strength lies in accessibility and ease of use, making it ideal for article authors rather than offering full geospatial analysis capabilities. Overall, our work does not aim to completely replace traditional methods of article writing and illustration creation or professional software, but rather to provide a more convenient and immersive approach.

Future Work. There are several directions that can be further explored. First, we can leverage modern text-to-image generative models to generate reference illustrations and recommend specific visual designs to enhance creativity. The system could also incorporate color recommendation or AI-assisted coloring features during the user refinement stage to address potential rigidity issues with initial colors generated during visualization creation, thereby improving visual appeal. Second, since GeoAuthor may generate geographic text with factual errors due to LLMs' limited geographic expertise, future work will fine-tune LLMs using geography-specific datasets to improve text generation quality. Last, we can consider adding a natural language interface to assist various features in GeoAuthor, including visualization creation, refinement, and text generation, through multiple rounds of human-machine dialogue instead of manual editing.

IX. CONCLUSION

Creating geographic articles with illustrations is common but often tedious and challenging. This difficulty stems from the lack of effective links between text and geographic visualizations, forcing authors to constantly switch between writing and creating illustrations, interrupting their workflow and immersive writing experience. We therefore developed GeoAuthor, which facilitates geographic article authoring by linking text and geographic visualizations. Our evaluation demonstrates GeoAuthor's usability and effectiveness, making geographic article creation more convenient and efficient.

ACKNOWLEDGMENTS

This work was supported by NSFC (U22A2032, 62532011, 62402421), Key "Pioneer" R&D Projects of Zhejiang Province (2026C01027), Ningbo Yongjiang Talent Programme (2024A-399-G), and Guangdong Basic and Applied Basic Research Foundation (2025A1515010162).

REFERENCES

- [1] W. H. Levie and R. Lentz, "Effects of text illustrations: A review of research," *Ectj*, vol. 30, no. 4, pp. 195–232, 1982.
- [2] A. Ottley, A. Kaszowska, R. J. Crouser, and E. M. Peck, "The curious case of combining text and visualization," *EuroVis 2019-Short Papers*, 2019.
- [3] "Leaflet," 2026, <https://leafletjs.com/> [Last accessed on 16 January 2026].
- [4] "Arcgis," 2026, <https://www.esri.com/> [Last accessed on 16 January 2026].
- [5] S. Latif, Z. Zhou, Y. Kim, F. Beck, and N. W. Kim, "Kori: Interactive synthesis of text and charts in data documents," *IEEE Transactions on Visualization and Computer Graphics*, vol. 28, no. 1, pp. 184–194, 2021.
- [6] R. Eccles, T. Kapler, R. Harper, and W. Wright, "Stories in geotime," *Information Visualization*, vol. 7, no. 1, pp. 3–17, 2008.
- [7] T. Gao, J. R. Hullman, E. Adar, B. Hecht, and N. Diakopoulos, "Newsviews: an automated pipeline for creating custom geovisualizations for news," in *Proceedings of the SIGCHI conference on human factors in computing systems*, 2014, pp. 3005–3014.
- [8] Y.-S. Kim, J. Hullman, and M. Agrawala, "Generating personalized spatial analogies for distances and areas," in *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, 2016, pp. 38–48.
- [9] A. Gunturu, B. Pearman, K. Ihara, M. Faraji, B. Wang, R. H. Kazi, and R. Suzuki, "Mapstory: Llm-powered text-driven map animation prototyping with human-in-the-loop editing," *arXiv preprint arXiv:2505.21966*, 2025.
- [10] H. Shao, H. Xu, and X. Yuan, "Goma: A grammar of semantic annotations on maps."
- [11] S. Latif, S. Chen, and F. Beck, "A deeper understanding of visualization-text interplay in geographic data-driven stories," in *Computer graphics forum*, vol. 40, no. 3. Wiley Online Library, 2021, pp. 311–322.
- [12] F. Beck and D. Weiskopf, "Word-sized graphics for scientific texts," *IEEE transactions on visualization and computer graphics*, vol. 23, no. 6, pp. 1576–1587, 2017.
- [13] Z. Chen and H. Xia, "Crossdata: Leveraging text-data connections for authoring data documents," in *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, 2022, pp. 1–15.
- [14] D. H. Kim, E. Hoque, J. Kim, and M. Agrawala, "Facilitating document reading by linking text and tables," in *Proceedings of the 31st Annual ACM Symposium on User Interface Software and Technology*, 2018, pp. 423–434.
- [15] S. K. Badam, Z. Liu, and N. Elmqvist, "Elastic documents: Coupling text and tables through contextual visualizations for enhanced document reading," *IEEE transactions on visualization and computer graphics*, vol. 25, no. 1, pp. 661–671, 2018.
- [16] T. Zhou, J. Huang, and G. Y.-Y. Chan, "Epigraphics: Message-driven infographics authoring," in *Proceedings of the CHI Conference on Human Factors in Computing Systems*, 2024, pp. 1–18.

- [17] A. Truong, F. Berthouzoz, W. Li, and M. Agrawala, "Quickcut: An interactive tool for editing narrated video," in *Proceedings of the 29th Annual Symposium on User Interface Software and Technology*, 2016, pp. 497–507.
- [18] S. Rubin, F. Berthouzoz, G. Mysore, W. Li, and M. Agrawala, "Under-score: musical underlays for audio stories," in *Proceedings of the 25th annual ACM symposium on User interface software and technology*, 2012, pp. 359–366.
- [19] B. Coyne and R. Sproat, "Wordseye: An automatic text-to-scene conversion system," in *Proceedings of the 28th annual conference on Computer graphics and interactive techniques*, 2001, pp. 487–496.
- [20] A. Chang, M. Savva, and C. D. Manning, "Learning spatial knowledge for text to 3d scene generation," in *Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP)*, 2014, pp. 2028–2038.
- [21] H. Xia, "Crosspower: Bridging graphics and linguistics," in *Proceedings of the 33rd Annual ACM Symposium on User Interface Software and Technology*, 2020, pp. 722–734.
- [22] W. Cui, X. Zhang, Y. Wang, H. Huang, B. Chen, L. Fang, H. Zhang, J.-G. Lou, and D. Zhang, "Text-to-viz: Automatic generation of infographics from proportion-related natural language statements," *IEEE transactions on visualization and computer graphics*, vol. 26, no. 1, pp. 906–916, 2019.
- [23] S. Latif, D. Liu, and F. Beck, "Exploring interactive linking between text and visualization," in *EuroVis (Short Papers)*, 2018, pp. 91–94.
- [24] N. Kong, M. A. Hearst, and M. Agrawala, "Extracting references between text and charts via crowdsourcing," in *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*, 2014, pp. 31–40.
- [25] N. Sultanum, F. Chevalier, Z. Bylinskii, and Z. Liu, "Leveraging text-chart links to support authoring of data-driven articles with vizflow," in *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, 2021, pp. 1–17.
- [26] D. H. Kim, S. Choi, J. Kim, V. Setlur, and M. Agrawala, "Emphasischecker: A tool for guiding chart and caption emphasis," *IEEE Transactions on Visualization and Computer Graphics*, vol. 30, no. 1, pp. 120–130, 2024.
- [27] J. Pinheiro and J. Poco, "Charttext: Linking text with charts in documents," *arXiv preprint arXiv:2201.05043*, 2022.
- [28] X. Cai, D. Weng, T. Fu, S. Fu, Y. Wang, and Y. Wu, "Linking text and visualizations via contextual knowledge graph," *IEEE Transactions on Visualization and Computer Graphics*, vol. 31, no. 9, pp. 4667–4681, 2025.
- [29] S. Latif and F. Beck, "Interactive map reports summarizing bivariate geographic data," *Visual Informatics*, vol. 3, no. 1, pp. 27–37, 2019.
- [30] C. Liu, L. Xie, Y. Han, D. Wei, and X. Yuan, "Autocaption: An approach to generate natural language description from visualization automatically," in *2020 IEEE Pacific visualization symposium (PacificVis)*. IEEE, 2020, pp. 191–195.
- [31] C. Lai, Z. Lin, R. Jiang, Y. Han, C. Liu, and X. Yuan, "Automatic annotation synchronizing with textual description for visualization," in *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, 2020, pp. 1–13.
- [32] K. Kafle, B. Price, S. Cohen, and C. Kanan, "Dvqa: Understanding data visualizations via question answering," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 5648–5656.
- [33] D. H. Kim, E. Hoque, and M. Agrawala, "Answering questions about charts and generating visual explanations," in *Proceedings of the 2020 CHI conference on human factors in computing systems*, 2020, pp. 1–13.
- [34] T. Marwah, G. Mittal, and V. N. Balasubramanian, "Attentive semantic video generation using captions," in *Proceedings of the IEEE international conference on computer vision*, 2017, pp. 1426–1434.
- [35] K. Chen, C. B. Choy, M. Savva, A. X. Chang, T. Funkhouser, and S. Savarese, "Text2shape: Generating shapes from natural language by learning joint embeddings," in *Computer Vision—ACCV 2018: 14th Asian Conference on Computer Vision, Perth, Australia, December 2–6, 2018, Revised Selected Papers, Part III 14*. Springer, 2019, pp. 100–116.
- [36] "Tableau," 2026, <https://www.tableau.com/trial/tableau-software/> [Last accessed on 16 January 2026].
- [37] "kepler.gl," 2025, <https://www.kepler.gl/> [Last accessed on 16 January 2026].
- [38] "Qgis," 2026, <https://qgis.org/> [Last accessed on 16 January 2026].
- [39] "Mapbox," 2026, <https://www.mapbox.com/> [Last accessed on 16 January 2026].
- [40] "deck.gl," 2026, <https://deck.gl/> [Last accessed on 16 January 2026].
- [41] J. Hynek, J. Kachlík, and V. Rusnák, "Geovisto: A toolkit for generic geospatial data visualization," in *VISIGRAPP (3: IVAPP)*, 2021, pp. 101–111.
- [42] J. Hynek and V. Rusnák, "Towards interactive geovisualization authoring toolkit for industry use cases," in *International Joint Conference on Computer Vision, Imaging and Computer Graphics*. Springer, 2021, pp. 232–256.
- [43] G. Moreira, M. Hosseini, M. N. Alam Nipu, M. Lage, N. Ferreira, and F. Miranda, "The urban toolkit: A grammar-based framework for urban visual analytics," *IEEE Transactions on Visualization and Computer Graphics*, vol. 30, no. 1, pp. 1402–1412, 2024.
- [44] C. Veiga, P. Li, F. Miranda, G. Moreira, S. Wu, K. Jogi, A. Gautam, M. Colombo, S. Tokkar, E. Makra *et al.*, "e-just-environmental justice using urban scalable toolkit: Development and enhancements," in *105th Annual AMS Meeting 2025*, vol. 105, 2025, p. 457659.
- [45] L. Ferreira, G. Moreira, M. Hosseini, M. Lage, N. Ferreira, and F. Miranda, "Assessing the landscape of toolkits, frameworks, and authoring tools for urban visual analytics systems," *Computers & Graphics*, vol. 123, p. 104013, 2024.
- [46] A. Satyanarayan, D. Moritz, K. Wongsuphasawat, and J. Heer, "Vegalite: A grammar of interactive graphics," *IEEE transactions on visualization and computer graphics*, vol. 23, no. 1, pp. 341–350, 2016.
- [47] Z. Deng, D. Weng, S. Liu, Y. Tian, M. Xu, and Y. Wu, "A survey of urban visual analytics: Advances and future directions," *Computational Visual Media*, vol. 9, no. 1, pp. 3–39, 2023.
- [48] M. Lin, Z. Deng, Q. Huang, Y. Ma, L.-P. Yuan, J. Bao, Y. Zheng, and Y. Cai, "Drainscope: Visual analytics of urban drainage system," *IEEE Transactions on Visualization and Computer Graphics*, 2026.
- [49] H. Chen, S. Jiang, X. Yu, H. Yin, X. Wang, Y. Hu, C. Wang, and C. Li, "Geovis: a data-driven geographic visualization recommendation system via latent space encoding: H. chen *et al.*" *Journal of Visualization*, vol. 27, no. 4, pp. 603–622, 2024.
- [50] F. Zheng, J. Wen, X. Zhang, Y. Chen, X. Zhang, Y. Liu, T. Xu, X. Chen, Y. Wang, W. Su *et al.*, "Visual abstraction of large-scale geographical point data with credible spatial interpolation," *Journal of Visualization*, vol. 24, no. 6, pp. 1303–1317, 2021.
- [51] C. Li, X. Dong, and X. Yuan, "Metro-wordle: An interactive visualization for urban text distributions based on wordle," *Visual Informatics*, vol. 2, no. 1, pp. 50–59, 2018.
- [52] H. Wang, Y. Ni, L. Sun, Y. Chen, T. Xu, X. Chen, W. Su, and Z. Zhou, "Hierarchical visualization of geographical areal data with spatial attribute association," *Visual Informatics*, vol. 5, no. 3, pp. 82–91, 2021.
- [53] H. Liu, X. Chen, Y. Wang, B. Zhang, Y. Chen, Y. Zhao, and F. Zhou, "Visualization and visual analysis of vessel trajectory data: A survey," *Visual Informatics*, vol. 5, no. 4, pp. 1–10, 2021.
- [54] Z. Deng, H. Chen, Q.-L. Lu, Z. Su, T. Schreck, J. Bao, and Y. Cai, "Visual comparative analytics of multimodal transportation," *Visual Informatics*, vol. 9, no. 1, pp. 18–30, 2025.
- [55] Z. Deng, S. Chen, X. Xie *et al.*, "Multilevel visual analysis of aggregate geo-networks [j/ol]," *IEEE Transactions on Visualization and Computer Graphics*, pp. 1–16, 2023.
- [56] Z. Deng, Y. Liu, M. Zhu, D. Xiang, H. Yu, Z. Su, Q. Lu, T. Schreck, and Y. Cai, "Trasculptor: Visual analytics for enhanced decision-making in road traffic planning," *IEEE Transactions on Visualization and Computer Graphics*, vol. 31, no. 10, pp. 6899–6914, 2025.
- [57] M. Wohlfart, "Story telling aspects in medical applications," in *Central European Seminar on Computer Graphics*, 2006.
- [58] M. Wohlfart and H. Hauser, "Story telling for presentation in volume visualization," in *Proceedings of the 9th Joint Eurographics/IEEE VGTC conference on Visualization*, 2007, pp. 91–98.
- [59] E. M. Lidal, M. Natali, D. Patel, H. Hauser, and I. Viola, "Geological storytelling," *Computers & graphics*, vol. 37, no. 5, pp. 445–459, 2013.
- [60] E. M. Lidal, H. Hauser, and I. Viola, "Geological storytelling: graphically exploring and communicating geological sketches," in *Proceedings of the international symposium on sketch-based interfaces and modeling*, 2012, pp. 11–20.
- [61] B. Lee, R. H. Kazi, and G. Smith, "Sketchstory: Telling more engaging stories with data through freeform sketching," *IEEE transactions on visualization and computer graphics*, vol. 19, no. 12, pp. 2416–2425, 2013.
- [62] P. Lundblad and M. Jern, "Geovisual analytics and storytelling using html5," in *2013 17th International Conference on Information Visualization*. IEEE, 2013, pp. 263–271.
- [63] J. Hullman, N. Diakopoulos, and E. Adar, "Contextifier: automatic generation of annotated stock visualizations," in *Proceedings of the*

SIGCHI Conference on human factors in computing systems, 2013, pp. 2707–2716.

- [64] J. Hullman, S. Drucker, N. H. Riche, B. Lee, D. Fisher, and E. Adar, “A deeper understanding of sequence in narrative visualization,” *IEEE Transactions on visualization and computer graphics*, vol. 19, no. 12, pp. 2406–2415, 2013.
- [65] F. Amini, N. Henry Riche, B. Lee, C. Hurter, and P. Irani, “Understanding data videos: Looking at narrative visualization through the cinematography lens,” in *Proceedings of the 33rd Annual ACM conference on human factors in computing systems*, 2015, pp. 1459–1468.
- [66] B. Bach, N. Kerracher, K. W. Hall, S. Carpendale, J. Kennedy, and N. Henry Riche, “Telling stories about dynamic networks with graph comics,” in *Proceedings of the 2016 CHI conference on human factors in computing systems*, 2016, pp. 3670–3682.
- [67] R. Swanson and A. S. Gordon, “Say anything: Using textual case-based reasoning to enable open-domain interactive storytelling,” *ACM Transactions on Interactive Intelligent Systems (TiIS)*, vol. 2, no. 3, pp. 1–35, 2012.
- [68] K. I. Gero and L. B. Chilton, “Metaphoria: An algorithmic companion for metaphor creation,” in *Proceedings of the 2019 CHI conference on human factors in computing systems*, 2019, pp. 1–12.
- [69] A. Coenen, L. Davis, D. Ippolito, E. Reif, and A. Yuan, “Wordcraft: A human-ai collaborative editor for story writing,” *arXiv preprint arXiv:2107.07430*, 2021.
- [70] J. J. Y. Chung, W. Kim, K. M. Yoo, H. Lee, E. Adar, and M. Chang, “Talebrush: Sketching stories with generative pretrained language models,” in *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, 2022, pp. 1–19.
- [71] R. Kiros, Y. Zhu, R. Salakhutdinov, R. S. Zemel, A. Torralba, R. Urtasun, and S. Fidler, “Skip-thought vectors,” in *Proceedings of the 29th International Conference on Neural Information Processing Systems - Volume 2*. MIT Press, 2015, pp. 3294–3302.
- [72] A. Fan, M. Lewis, and Y. Dauphin, “Hierarchical neural story generation,” *arXiv preprint arXiv:1805.04833*, 2018.
- [73] A. Khalifa, G. A. Barros, and J. Togelius, “Deeptingle,” *arXiv preprint arXiv:1705.03557*, 2017.
- [74] P. Ammanabrolu, E. Tien, W. Cheung, Z. Luo, W. Ma, L. J. Martin, and M. O. Riedl, “Story realization: Expanding plot events into sentences,” in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 34, no. 05, 2020, pp. 7375–7382.
- [75] V. Braun and V. Clarke, *Thematic analysis*. American Psychological Association, 2012.
- [76] “Rednote,” 2026, <https://www.xiaohongshu.com/> [Last accessed on 16 January 2026].
- [77] Y. Zheng, W. Wu, Y. Chen, H. Qu, and L. M. Ni, “Visual analytics in urban computing: An overview,” *IEEE Transactions on Big Data*, vol. 2, no. 3, pp. 276–296, 2016.
- [78] T. Brown, B. Mann, N. Ryder, M. Subbiah, J. D. Kaplan, P. Dhariwal, A. Neelakantan, P. Shyam, G. Sastry, A. Askell *et al.*, “Language models are few-shot learners,” *Advances in neural information processing systems*, vol. 33, pp. 1877–1901, 2020.
- [79] j. Brooke, “SUS: A ‘Quick and Dirty’ Usability Scale,” in *Usability Evaluation In Industry*. CRC Press, 1996.
- [80] J. R. Lewis and J. Sauro, “The Factor Structure of the System Usability Scale,” in *Proceedings of the Human Centered Design*, 2009, pp. 94–103.



Zhenning Chen received his B.S. degree in Software Engineering from Harbin Institute of Technology in 2023. He is currently a Ph.D. student in the State Key Lab of CAD&CG, Zhejiang University. His main research interest lies in information visualization and visual analytics, particularly geographic visualization and visual analytics for LLMs.



Hanbei Zhan received her B.S. degree in Computer Science and Technology from Zhejiang University in 2025. She is currently a master student at the State Key Lab of CAD&CG, Zhejiang University. Her research focuses on geographic visualization and visual analytics for LLMs.



Shifu Chen received his B.S. degree in Software Engineering from University of Electronic Science and Technology of China in 2022. He is currently a Ph.D. student in the State Key Lab of CAD&CG, Zhejiang University. His research interests mainly include visualization and visual analytics.



Zikun Deng is a tenure-track associate professor at School of Software Engineering, South China University of Technology. He received his Ph.D. degree in Computer Science from the State Key Lab of CAD&CG, Zhejiang University in 2023. His research interests mainly include visual analytics, visualization, data mining, and their application in smart city, industry 4.0, and digital twins. He has published more than 10 papers in IEEE TVCG. For more information, please visit <https://zkdeng.org>.



Di Weng is a ZJU100 Young Professor at the School of Software Technology, Zhejiang University. His main research interest lies in information visualization and visual analytics, focusing on interactive data transformation and spatiotemporal data analysis. He received his Ph.D. degree in Computer Science from the State Key Lab of CAD&CG, Zhejiang University. Before his current position, Dr. Weng was a researcher at Microsoft Research Asia from 2022 to 2023. For more information, please visit <https://dwe.ng>.



Yingcai Wu is a Professor at the State Key Lab of CAD&CG, Zhejiang University. His main research interests are information visualization and visual analytics, with focuses on urban computing, sports science, immersive visualization, and social media analysis. He received his Ph.D. degree in Computer Science from The Hong Kong University of Science and Technology. Prior to his current position, Dr. Wu was a postdoctoral researcher in the University of California, Davis, from 2010 to 2012, and a researcher at Microsoft Research Asia from 2012 to 2015. For more information, please visit <http://www.ycwu.org>.