A Demo of GeoQA^3: Towards An Accessible AI-based Question-Answering System for Geoanalytics

Chu Li
Paul G. Allen School of Computer Science and Engineering
University of Washington
Seattle, Washington, USA
chuchuli@cs.washington.edu

Rock Yuren Pang
Paul G. Allen School of Computer Science and Engineering
University of Washington
Seattle, Washington, USA
ypang2@cs.washington.edu

Arnavi Chheda-Kothary
Paul G. Allen School of Computer Science and Engineering
University of Washington
Seattle, Washington, USA
chheda@cs.washington.edu

Ather Sharif
Paul G. Allen School of Computer Science and Engineering
University of Washington
Seattle, Washington, USA
asharif@cs.washington.edu

Henok Assalif
Paul G. Allen School of Computer Science and Engineering
University of Washington
Seattle, Washington, USA
henok206@uw.edu

Jeffrey Heer
Paul G. Allen School of Computer Science and Engineering
University of Washington
Seattle, Washington, USA
jheer@uw.edu

Jon E. Froehlich
Paul G. Allen School of Computer Science and Engineering
University of Washington
Seattle, Washington, USA
jonf@cs.uw.edu

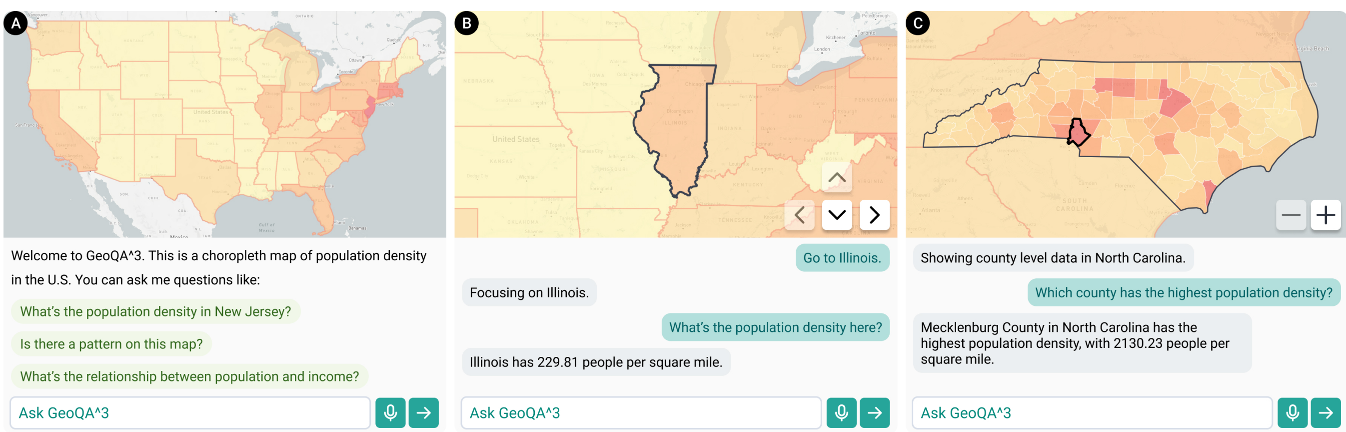

Figure 1: We introduce GeoQA^3, a novel accessible AI-based question-answering system for geovisualizations designed for screen-reader users. (A) Through a custom query pipeline, we combine geo-statistical analysis with an LLM to balance accuracy and performance. (B) Users can navigate the map through natural language commands or keyboard controls and (C) zoom in to view county-level data. The AI Chat system is context-aware, taking into account user interactions. See video for demonstration.

Abstract

Geovisualizations are powerful tools for analyzing and interpreting spatial data; however, they are historically inaccessible to screen-reader users. We introduce GeoQA^3, an **A**ccessible **AI**-based **Q**uestion-**A**nswering (QA) system to enable blind users to perform geo-spatial analytics. GeoQA^3 relies on a custom QA pipeline that combines map interactions with chat questions to form queries and uniquely combines geo-statistical analysis with LLM-based summaries. In a remote lab study with six screen-reader users, we found that participants successfully employed diverse querying strategies for spatial analysis and highly valued the AI Chat component for its interactive responses. During the ASSETS demo session, attendees will use GeoQA^3 to explore two key questions, mirroring our user

ASSETS '25, Denver, CO, USA

ACM ISBN 979-8-4007-0676-9/25/10
https://doi.org/10.1145/3663547.3759700

study: potential biases in digital access in the US and how energy sources differ geographically across the US.

Keywords

Accessible visualizations, geoanalytics, visualization question and answering

ACM Reference Format:

Chu Li, Rock Yuren Pang, Arnavi Chheda-Kothary, Ather Sharif, Henok Assalif, Jeffrey Heer, and Jon E. Froehlich. 2025. A Demo of GeoQA^3: Towards An Accessible AI-based Question-Answering System for Geoanalytics. In *The 27th International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '25), October 26–29, 2025, Denver, CO, USA.* ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/3663547.3759700

1 Introduction

Interactive map visualizations, or *geovisualizations*, are powerful tools that allow users to discern and analyze patterns, trends, and relationships in spatial data [1, 23]. Despite their importance, most geovisualizations are inaccessible to screen-reader users due to their inherent reliance on spatial data and visual form [6, 21, 22]. Even when accessibility features such as *alt text* and *data tables* are available, they do not capture the full analytical and interpretive potential of geovisualizations [8, 11].

While recent work like *AltGeoViz* [12] offer dynamic alt text for geovisualizations based on user selections and zoom, these approaches remain largely descriptive, akin to "map reading" rather than "map analysis" [13]. Our goal is to support deeper spatial analysis—allowing screen reader users to identify patterns, relationships, and geometric characteristics that are crucial for insight building and interpretation [15, 16]. Emerging question-answering (QA) systems for geovisualizations such as *MapQA* [4] and *VoxLens* [19–21] are an important step but rely on keyword matching or rigid rule sets, limiting interactive queries.

In this demo paper, we introduce GeoQA^3, an accessible AI-based question-answering system for geoanalytics. Through a custom interactive query pipeline that combines geo-statistical analysis with an LLM, GeoQA^3 uniquely supports seven analytical query types for screen reader users, including: retrieve, compare, find extrema, sort, filter, compute derived values, and cluster by values [10]. Moreover, GeoQA^3 specifically handles geospatial queries for *spatial patterns* [18], *spatial relationships*, and *geometric characteristics.* To ease interaction, users can fluidly switch between keyboard-based navigation and conversational commands for map exploration. GeoQA^3 was designed iteratively using co-design with two screen-reader users and drawing on design principles from QA literature, including disambiguating deictic references [9, 10] and supporting general/contextual queries [8, 10]). See video demo.

To evaluate GeoQA^3, we conducted a user study with six screen-reader users asking participants to perform exploratory data analysis for two tasks: (1) distributing digital equity funding based on underserved populations and lack of digital access, and (2) identifying predominant energy sources in different regions and explaining potential reasons. We found that participants successfully employed diverse querying strategies for spatial analysis, valued the chat component for its informative and clear responses, and highlighted the system's navigation autonomy, while also identifying areas for improvement in navigation complexity and answer specificity.

In summary, we contribute: (1) GeoQA^3, a novel LLM-powered interactive QA system for accessible geoanalytics; (2) an initial validation with six screen-reader users demonstrating effectiveness and important areas for future work. During the ASSETS'25 demo session, attendees will be invited to use GeoQA^3 to analyze our two study datasets (digital equity, U.S. energy sources), interact with our custom QA pipeline, and discuss future work.

2 The GeoQA^3 Prototype

GeoQA^3 is composed of two primary components: (1) a screen-reader compatiable UI with an interactive map and an AI-based chat that supports analytical, geospatial, visual, and contextual queries; (2) a custom QA pipeline that combines map interactions with chat questions to form queries and uniquely combines geo-statistical analysis with LLM-based summaries. GeoQA^3's frontend is implemented in MapboxJS and React.js, and the backend in Python's Flask framework. For the LLM, we use GPT-4o-mini, which balances computational efficiency and performance [17]. We begin by describing the QA pipeline as it is a central technical contribution of our work.

2.1 QA Pipeline

Our custom QA pipeline consists of four components (Figure 2): *Input Classifier*, *Query Refiner*, *Scope Assessor*, and *Query Processor*. We used few-shot prompting with LLMs for all pipeline components (rather than fine-tuning) due to its high performance in traditional classification tasks [3, 7]. See Supplementary Materials for the prompts.

Input Classifier. Upon receiving user input, GeoQA^3 first determines whether to perform an *action command* or *information query*. Action commands, such as *"Pan to Minnesota"* or *"Zoom to Minneapolis"* trigger direct map manipulation while information queries proceed to Query Refiner for further processing.

Query Refiner. Because natural language queries can be ambiguous, the Query Refiner addresses (1) *location ambiguity*, where deictic references such as *"here"* or *"this state"* are resolved using the current map focus or previous conversation, *e.g., "What's the population density here?"* when a specific state is highlighted; and (2) *topic ambiguity*, where pronouns like *"that"* or *"it"* are resolved using conversation history, *e.g., "How does that compare to Ohio?"* where *"that"* refers to previously discussed population density. Some queries exhibit both ambiguity types, *e.g., "How does it compare to its neighbors?"* where *"it"* refers to a previously discussed metric and *"its"* refers to the focused state.

Scope Assessor. Following disambiguation, GeoQA^3 evaluates whether the query falls within the scope of local processing capabilities: *within-scope queries* proceed to the Query Processor for classification and local operations, while *beyond-scope* ones are routed to GPT-4o-mini. For example, queries requesting Idaho's population density are resolved using local datasets when such data is available, whereas queries about Idaho's median household income are routed to GPT-4o-mini when income data is not present in our local repository.

Query Processor. The Query Processor classifies the query type according to Kim *et al.'s* visualization query taxonomy [10] (extended for geovisualization interactions): *map actions*, *analytical queries* such as retrieval, comparison, or aggregation; *geospatial queries* such as pattern or outlier detection; *visual queries* such as color, shape, or spatial relationship queries; *contextual queries*

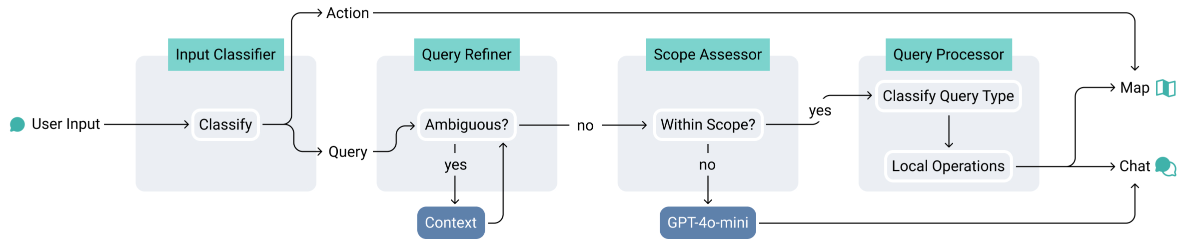

Figure 2: The QA pipeline consists of four components: Input Classifier, Query Refiner, Scope Assessor, and Query Processor and uniquely combines geostatistical analyses with LLM output to produce accurate calculations with explanatory summaries.

such as visualization concept queries (*e.g.,* "What is a choropleth map?") or general knowledge queries (*e.g.,* "Is there a relationship between income and population density?"). Crucially, for both analytical and geospatial queries, we conduct analyses locally without the LLM—this ensures both valid and efficient responses. For analytical queries, GeoQA^3 executes appropriate local operations on the dataset such as sorting, clustering, comparison. For geospatial queries, GeoQA^3 conducts spatial statistical analysis by (1) computing global Moran's I [14] to identify overall spatial autocorrelation patterns, then (2) calculating Local Indicators of Spatial Association (LISA) clusters [2]. Rather than showing raw statistical output, we use GPT-4o-mini to produce interpretable summary explanations. Queries that fall outside these types are processed by GPT-4o-mini.

Output Generation. Finally, we output results through coordinated updates to both the map and AI chat. Geospatial analyses are visualized as an additional map layer showing LISA clusters represented as colored outlines on the map, providing visual confirmation of the patterns described in the textual response.

2.2 User Interface

GeoQA^3's UI is composed of a map visualization and an AI Chat subsystem—both feed into the QA pipeline. Screen readers parse the map and messages and convey them through audio to participants.

Interactive Map Though customizable for other regions, our current map visualization includes state- and county-level representations. To start map interactions, users can press Ctrl + M and then Tab to focus on a state as well as arrow keys to move between states. Since the four cardinal directions are insufficient to represent all possible spatial relationships (*e.g.,* both Ohio and Indiana are north of Kentucky), we developed a custom algorithm to select at most one neighboring state for each cardinal direction. The algorithm: (1) identifies adjacent states by detecting shared boundaries, (2) determines cardinal directions between states by comparing their centroids, and (3) selects the closest neighbor in each cardinal direction for each state. For states with complex geometries and atypical centroid positions, we implemented manual adjustments to ensure natural navigation (*e.g.,* New York, DC, Rhode Island). When users attempt to navigate in a direction where no neighboring state exists, the system indicates the boundary condition, *e.g., "There is no state south of Texas"*. Once focused on a state, users can press + to zoom in to county-level data. The system automatically focuses on the county closest to the state's centroid, after which similar arrow key navigation becomes available at the county level. Users can return to the state-level view by pressing the − key.

AI Chat The AI chat interface follows conversational UI design standards while accommodating screen-reader interaction. Users can toggle focus between map and chat by using Ctrl + M. Upon focus, GeoQA^3 announces, *"Type your question here, press enter to submit."*. Following question submission, the system repeats the user query followed by a status indicator (*"Looking for answers..."*). Users can navigate to previous conversation history by using Tab. Ctrl + L repeats the most recent system response while maintaining the input field focus.

When AI Chat first loads, we provide an introduction to GeoQA^3 and provide three selectable example geoanalytic questions. Below these examples, a *More Suggestions* button (accessible via Ctrl + I) refreshes the question set, cycling through a list of 12 predefined example questions. AI Chat also displays contextual questions that extend beyond the immediate dataset, *e.g., "What is a choropleth map?"*. For additional assistance, users can ask *"What else can you do?"* to retrieve supported query types. The Ctrl + H shortcut displays a comprehensive list of navigation commands.

Map and Chat Synchronization To support a tightly integrated, holistic interactive experience, the map and chat components are bidirectionally synchronized. When users query specific geographic entities through AI Chat, the system automatically updates the map to provide relevant spatial context. For example, when a user asks about a single state's value (*e.g., "What is the population density of Illinois?"*), the map centers on Illinois and highlights the boundary (Figure 1B). Beyond queries, GeoQA^3 also responds to explicit navigation commands in natural language, such as *"Take me to Wyoming"*, *"Focus on Cook County, Illinois"*, or *"Go to Sacramento"*, by immediately updating the map focus accordingly. Moreover, GeoQA^3 maintains contextual awareness of the user's current map focus during free exploration, enabling implicit geographic referencing in AI Chat. For example, when focused on Colorado, users can ask *"What's the population density here?"* or *"What are the neighboring states?"*.

3 User Study

To evaluate GeoQA^3, we conducted a 90-120 min remote user study with six screen-reader users. Participants were completely blind (ages 25-64), used JAWS or NVDA screen readers, and had limited geovisualization experience. As the first AI-based accessible geoanalytic system, our study goals were twofold: first, to explore what *types of queries* do screen reader users make when interacting with an accessible geovisualization; and second, to examine *how well* the current GeoQA^3 prototype supports these queries, builds appropriate mental models of the underlying data, and leads to accurate insights and takeaways. All study sessions were audio and video recorded and we used a combination of measures to address our research goals.

Study Tasks. After a brief tutorial, participants were asked to perform two geoanalytical tasks. For both tasks, we emphasized interest in studying participants' analytical approaches and not their derived answers. In Task 1, participants were asked to imagine themselves as decision-makers responsible for distributing *State Digital Equity Planning Grant* funding across states—a task adapted from the U.S. Census on mapping digital equity, which originally contained inaccessible geovisualizations[1]. Participants were told to select both a single state and a cluster of four to six geographically contiguous states most deserving of funding allocation. For Task 2, we adapted a *Washington Post* news article entitled "*U.S. Home Heating is Fractured in Surprising Ways*"[2] into GeoQA^3. Participants were asked to identify predominant heating fuel sources in different U.S. regions and consider potential explanations for these patterns.

Findings. Overall, participants found GeoQA^3 to be a valuable tool for engaging with geovisualizations, demonstrating map reading, analysis, interpretation, and navigation. They unanimously rated the chat component as 7/7, highlighting its effectiveness in providing information. Participants employed diverse query strategies, including ranking, sorting, comparing, and inquiring about geographic patterns and spatial relationships, often relying on a combination of verbal queries and key navigation. The system effectively supported basic map reading and facilitated the recall of geographic knowledge, with participants describing responses as *"very informative"*, *"clear and concise"*, and *"very accurate"*. GeoQA^3 also enabled BLV users to interpret complex spatial data, facilitating a deeper understanding of patterns and relationships and allowing them to query across datasets for more specific insights.

While promising, participants also identified important areas for future work. They requested enhanced navigation beyond cardinal directions and more specific responses with source attribution. Moreover, they suggested future domains of interest including the interpretation of election maps, accessing news-related information, and understanding resource distribution, with half of them requesting continued access to the system post-study.

4 Discussion and Conclusion

In this demo paper, we introduced GeoQA^3, a novel accessible AI-based question-answering system for geovisualizations designed for screen-reader users. Our study revealed key insights into how LLM-based QA systems can support screen-reader users in exploring and analyzing geovisualizations. While GeoQA^3 demonstrates potential, we also observed areas for improvement. For instance, participants sometimes misinterpreted spatial patterns, such as incorrectly identifying states with the highest values, necessitating additional queries about spatial relationships to confirm connections. Furthermore, traditional cluster identification methods like Moran's I have inherent limitations in fully representing patterns that might be more intuitively understood through other means. Future work should explore advanced pattern detection approaches such as incorporating estimate errors.

While participants generally trusted GeoQA^3's answers, they lacked straightforward methods to verify responses. This underscores the importance of communicating uncertainty and providing source attribution for AI-generated information [5]. Additionally, our study revealed that successful interaction often required a degree of data literacy from participants, who sometimes struggled with query formulation (*e.g.,* asking about "fuel" instead of "heating fuel"), which returned results about automotive fuel rather than household energy. This highlights a clear need for dynamic guided prompting, where GeoQA^3 proactively clarifies user intent during the querying process.

Finally, screen-reader participants expressed a strong desire for more direct manipulation capabilities, such as the ability to select multiple states simultaneously for comparison and analysis. This interest reflects a broader challenge inherent in purely language-based interfaces.

Acknowledgments

This work was supported by NSF SCC-IRG #2125087.

References

[1] Natalia Andrienko and Gennady Andrienko. 2006. *Exploratory analysis of spatial and temporal data: a systematic approach.* Springer Science & Business Media.

[2] Luc Anselin. 1995. Local Indicators of Spatial Association—LISA. *Geographical Analysis* 27, 2 (April 1995), 93–115. doi:10.1111/j.1538-4632.1995.tb00338.x

[3] Tom Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared D. Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, and Amanda Askell. 2020. Language models are few-shot learners. *Advances in neural information processing systems* 33 (2020), 1877–1901. https://proceedings.neurips.cc/paper/2020/hash/1457c0d6bfcb4967418bfb8ac142f64a-Abstract.html

[4] Shuaichen Chang, David Palzer, Jialin Li, Eric Fosler-Lussier, and Ningchuan Xiao. 2022. MapQA: A Dataset for Question Answering on Choropleth Maps. doi:10.48550/arXiv.2211.08545 arXiv:2211.08545.

[5] Hyo Jin Do, Rachel Ostrand, Justin D. Weisz, Casey Dugan, Prasanna Sattigeri, Dennis Wei, Keerthiram Murugesan, and Werner Geyer. 2024. Facilitating Human-LLM Collaboration through Factuality Scores and Source Attributions. *ArXiv* abs/2405.20434 (2024). https://api.semanticscholar.org/CorpusID:270199456

[6] Danyang Fan, Alexa Fay Siu, Hrishikesh Rao, Gene Sung-Ho Kim, Xavier Vazquez, Lucy Greco, Sile O'Modhrain, and Sean Follmer. 2023. The Accessibility of Data Visualizations on the Web for Screen Reader Users: Practices and Experiences During COVID-19. *ACM Transactions on Accessible Computing* 16, 1 (March 2023), 1–29. doi:10.1145/3557899

[7] Tianyu Gao, Adam Fisch, and Danqi Chen. 2021. Making Pre-trained Language Models Better Few-shot Learners. doi:10.48550/arXiv.2012.15723 arXiv:2012.15723 [cs].

[8] Joshua Gorniak, Yoon Kim, Donglai Wei, and Nam Wook Kim. 2024. VizAbility: Enhancing Chart Accessibility with LLM-based Conversational Interaction. In *Proceedings of the 37th Annual ACM Symposium on User Interface Software and Technology.* ACM, Pittsburgh PA USA, 1–19. doi:10.1145/3654777.3676414

[9] Enamul Hoque, Vidya Setlur, Melanie Tory, and Isaac Dykeman. 2018. Applying Pragmatics Principles for Interaction with Visual Analytics. *IEEE Transactions on Visualization and Computer Graphics* 24, 1 (Jan. 2018), 309–318. doi:10.1109/TVCG.2017.2744684 Conference Name: IEEE Transactions on Visualization and Computer Graphics.

[10] Jiho Kim, Arjun Srinivasan, Nam Wook Kim, and Yea-Seul Kim. 2023. Exploring Chart Question Answering for Blind and Low Vision Users. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (CHI '23).* Association for Computing Machinery, New York, NY, USA, 1–15. doi:10.1145/3544548.3581532

[11] N. W. Kim, G. Ataguba, S. C. Joyner, Chuangdian Zhao, and Hyejin Im. 2023. Beyond Alternative Text and tables: Comparative Analysis of Visualization Tools and Accessibility Methods. *Computer Graphics Forum* 42, 3 (2023), 323–335. doi:10.1111/cgf.14833 _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/cgf.14833.

[1] https://www.census.gov/library/stories/2022/05/mapping-digital-equity-in-every-state.html

[2] https://www.washingtonpost.com/climate-environment/interactive/2023/home-electrification-heat-pumps-gas-furnace/

[12] Chu Li, Rock Yuren Pang, Ather Sharif, Arnavi Chheda-Kothary, Jeffrey Heer, and Jon E. Froehlich. 2024. AltGeoViz: Facilitating Accessible Geovisualization. In *2024 IEEE Visualization and Visual Analytics (VIS)*. 61–65. doi:10.1109/VIS55277.2024.00020 ISSN: 2771-9553.

[13] Janet E. Mersy. 1990. Choropleth map design - a map user study. *Cartographica* 27, 3 (Sept. 1990), 33–50. doi:10.3138/1928-QQ57-3625-L024 Publisher: University of Toronto Press.

[14] Patrick AP Moran. 1948. The interpretation of statistical maps. *Journal of the Royal Statistical Society. Series B (Methodological)* 10, 2 (1948), 243–251. https://www.jstor.org/stable/2983777 Publisher: JSTOR.

[15] Phillip Muehrcke. 1978. Functional map use. *Journal of Geography* 77, 7 (Dec. 1978), 254–262. https://www.proquest.com/docview/1290633193/citation/D967AF4B20D247BCPQ/1 Num Pages: 9 Place: Macomb, Ill., etc., United States Publisher: National Council for Geographic Education.

[16] Judy M. Olson. 1976. A Coordinated Approach to Map Communication Improvement. *The American Cartographer* 3, 2 (Jan. 1976), 151–160. doi:10.1559/152304076784080177 Publisher: Taylor & Francis _eprint: https://doi.org/10.1559/152304076784080177.

[17] OpenAI. 2024. GPT-4o mini: advancing cost-efficient intelligence. https://openai.com/index/gpt-4o-mini-advancing-cost-efficient-intelligence/

[18] Jochen Schiewe. 2019. Empirical Studies on the Visual Perception of Spatial Patterns in Choropleth Maps. *KN - Journal of Cartography and Geographic Information* 69, 3 (Sept. 2019), 217–228. doi:10.1007/s42489-019-00026-y

[19] Ather Sharif, Olivia H. Wang, Alida T. Muongchan, Katharina Reinecke, and Jacob O. Wobbrock. 2022. VoxLens: Making Online Data Visualizations Accessible with an Interactive JavaScript Plug-In. In *CHI Conference on Human Factors in Computing Systems*. ACM, New Orleans LA USA, 1–19. doi:10.1145/3491102.3517431

[20] Ather Sharif, Andrew M. Zhang, Katharina Reinecke, and Jacob O. Wobbrock. 2023. Understanding and Improving Drilled-Down Information Extraction from Online Data Visualizations for Screen-Reader Users. In *20th International Web for All Conference*. ACM, Austin TX USA, 18–31. doi:10.1145/3587281.3587284

[21] Ather Sharif, Andrew Mingwei Zhang, Anna Shih, Jacob O. Wobbrock, and Katharina Reinecke. 2022. Understanding and Improving Information Extraction From Online Geospatial Data Visualizations for Screen-Reader Users. In *Proceedings of the 24th International ACM SIGACCESS Conference on Computers and Accessibility*. ACM, Athens Greece, 1–5. doi:10.1145/3517428.3550363

[22] Jonathan Zong, Crystal Lee, Alan Lundgard, JiWoong Jang, Daniel Hajas, and Arvind Satyanarayan. 2022. Rich Screen Reader Experiences for Accessible Data Visualization. *Computer Graphics Forum* 41, 3 (June 2022), 15–27. doi:10.1111/cgf.14519

[23] Arzu Çöltekin, Halldór Janetzko, Sara Fabrikant, University of Zurich, and University of Zurich. 2018. Geovisualization. *Geographic Information Science & Technology Body of Knowledge* 2018, Q2 (April 2018). https://gistbok-topics.ucgis.org/CV-05-035