VizAbility: Multimodal Accessible Data Visualization with Keyboard Navigation and Conversational Interaction

Joshua Gorniak
joshua.gorniak@bc.edu
Boston College
Chestnut Hill
Massachusetts, USA

Jake Ottiger
jake.ottiger@bc.edu
Boston College
Chestnut Hill
Massachusetts, USA

Donglai Wei
donglai.wei@bc.edu
Boston College
Chestnut Hill
Massachusetts, USA

Nam Wook Kim
nam.wook.kim@bcu
Boston College
Chestnut Hill
Massachusetts, USA

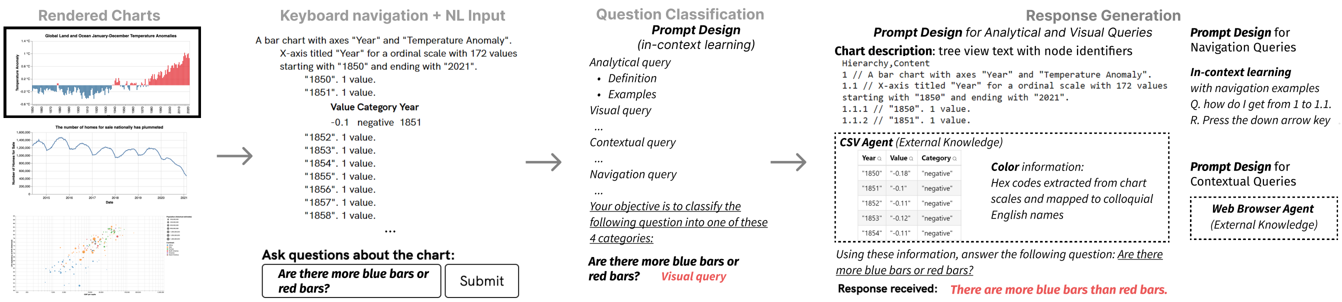

Figure 1: VizAbility pipeline: users navigate the chart using a keyboard and ask questions that are answered by classifying their query type (e.g., visual query) and referring to underlying data, chart visual structure, user location, and internet browsing.

ABSTRACT

Data visualization serves as a crucial tool for communicating important information in our society. Yet, as visualizations grow more complex, they become less accessible to individuals with visual impairments. Traditional accessibility approaches like alternative text and data tables often fall short of capturing the full potential of data visualization. To bridge this gap, we introduce VizAbility, a novel multimodal accessible system that combines keyboard navigation with conventional interaction, enabling individuals with visual impairments to actively engage with and explore data visualizations. We built an LLM-based pipeline that classifies user queries and synthesizes underlying data, chart structure, user locality, and web-based information to answer the queries. Our preliminary evaluation using real-world questions from blind individuals demonstrates the significant potential of VizAbility.

CCS CONCEPTS

• **Human-centered computing** → **Interactive systems and tools**; **Visualization systems and tools**.

KEYWORDS

data visualization, accessibility, blind and low vision people

UIST '23 Adjunct, October 29–November 01, 2023, San Francisco, CA, USA

ACM ISBN 979-8-4007-0096-5/23/10.
https://doi.org/10.1145/3586182.3616669

ACM Reference Format:
Joshua Gorniak, Jake Ottiger, Donglai Wei, and Nam Wook Kim. 2023. VizAbility: Multimodal Accessible Data Visualization with Keyboard Navigation and Conversational Interaction. In *The 36th Annual ACM Symposium on User Interface Software and Technology (UIST '23 Adjunct), October 29–November 01, 2023, San Francisco, CA, USA.* ACM, New York, NY, USA, 3 pages. https://doi.org/10.1145/3586182.3616669

1 INTRODUCTION

Data visualizations are widely used to convey complex data. However, their inaccessibility poses a significant challenge for those who are blind or have low vision, leading to an information disparity, especially with crucial data like health information [12, 14]. Regrettably, the majority of data visualizations found in practice remain inaccessible [5], and existing visualization tools do not support accessible design well [8]. While alternative text and data tables are commonly used methods for accessible visualizations, they tend to diminish the inherent benefits that data visualizations offer [9].

This paper presents VizAbility, a multimodal accessible data visualization system that integrates keyboard navigation and conversational interaction. VizAbility incorporates ideas from recent approaches, specifically keyboard-navigable chart content [16, 17] and chart question & answering [13]. Instead of solely focusing on an individual modality, we integrate the respective advantages of these approaches to provide an enhanced accessible data experience. Unlike previous multimodal accessible systems that require custom tactile or haptic hardware with additional audio feedback [3, 4, 9], our system eliminates any additional adoption cost for screen reader users and can be easily scaled for diverse chart types. In the following sections, we describe our design decisions and provide an overview of the system pipeline.

2 VIZABILITY DESIGN DECISIONS

Our work is motivated by a study that highlights the advantages of keyboard navigation for chart content and conversational interaction [8]. While keyboard navigation can be cumbersome when dealing with complex charts, it closely represents the visual layout. Conversely, conversational interaction allows efficient data exploration through natural language queries but lacks independent verification. These approaches complement each other, offering distinct benefits while mitigating limitations.

To create VizAbility, we built an LLM-based pipeline that handles different user inquiries, including analytical, visual, contextual, and navigational queries. This pipeline integrates Olli [2], an open-source library that transforms charts into a keyboard-navigable structure [17]. We use LLMs due to their broad knowledge base and language understanding capability, which significantly advances the existing system, VoxLens [13], which can only handle single-series data. We improve response quality and enable navigation queries by incorporating Olli information into our prompt design.

3 VIZABILITY SYSTEM PIPELINE

VizAbility assumes that visual encoding information and the underlying dataset are available in the Vega-lite [11] spec. We feed this spec to Olli [2] to render the chart content in a keyboard-navigable format. We also render the chart graphic for sighted users. The text box enables blind users to type their questions regarding the chart.
Step 1: Query Classification. Our system classifies users' queries into one of four classes: analytical query (e.g., involving only data), visual query (e.g., involving encoding info), contextual query, and navigation query. Such a task-division approach is known to improve LLM performance [15]. We use a few-shot prompting approach by providing a few examples of our own ground-truth classifications of the questions used in Kim *et al.* [7]. We explicitly instruct the LLM to output only the query type without any additional text. If the user query cannot be classified, VizAbility responds, *"I am sorry I am unable to answer the question"*.
Step 2: Query-specific Prompting. We then construct different prompts based on the query type. For ANALYTICAL QUERIES that primarily involve data, we utilize the CSV agent in LangChain [1], which leverages OpenAI's API (`gpt-3.5-turbo` with `temperature = 0`) [10]; this allows us to work around the token limit in the prompt by putting data as external knowledge in a CSV file. Since Vega-lite may transform the raw data to generate a chart, we send the final transformed view data to the CSV agent [1]. To further help contextualize user questions, we take advantage of Olli's tree view by sending its textual description of the chart and the user's location in the tree. For instance, for the analytic question *"What was the least amount of houses sold?"* for the line chart (Figure 1), VizAbility responds, *"The least amount of houses sold was 468,297."*

To handle VISUAL QUERIES that include visual encoding information, we also extract color encoding information from the scales in the parsed Vega-lite spec and include it in the CSV file along with the data. As the color encodings are in hex codes, we convert them into common English color names. The LLM can then determine the color the user is referring to by using the closest color name (`user: red → orangered`). For instance, when the question is "*What color is Africa?"* for the scatter plot (Figure 1), VizAbility responds with *"Africa is teal."*. The remaining pipeline follows the same architecture as that of analytical queries.

For CONTEXTUAL QUERIES that do not require the information depicted in the chart but ask about general relevant knowledge, we employ a Web Browser agent [1]. For instance, when asked the question *"What do we mean by temperature anomalies?"* for the bar chart (Figure 1), VizAbility responds, *"Temperature anomalies are deviations from a reference or average temperature..."*

On the other hand, NAVIGATION QUERIES pertain to users' location within the tree view. For instance, the user might ask *"How do I get to the X-axis from here"* for the scatter plot (Figure 1), VizAbility then guides the user by responding, *"Press the up arrow key. Press the left arrow key."*. To handle these navigation queries, we assign a unique address to each node in the tree view (Figure 1, e.g., `1.2` is the second child of the root. We send this, along with the user's active position within the tree view. Utilizing few-shot prompting once more using example navigation from one point to another, the LLM is trained to format its answer into a step-by-step guide.

4 PRELIMINARY EVALUATION

Dataset. We used a question dataset from prior work [7] for evaluation. The dataset consists of four charts (bar, line, scatter, and map) and 979 questions from blind individuals. We excluded the choropleth map as Olli [2] cannot handle it yet. We converted the charts into Vega-lite specs for testing and manually categorized them into the four query types. The questions did not include navigation queries. Next, we performed an 80/20 split between the testing and validation sets through stratified random sampling based on query classification; we used the validation set to construct examples for few-shot prompting. Finally, we randomly extracted 5% of the testing set, resulting in 33 questions for preliminary evaluation, considering computational time and cost constraints.
Results. For step 1, VizAbility achieved 91% (30/33) accuracy in categorizing user queries. For step 2, out of the 30 questions, it successfully answered 76.7% (23/30) when we manually assessed the responses as "correct" or "incorrect". When considering only bar and line charts, our accuracy is 60% (12/20). For comparison, the existing systems [6, 7] can only handle bar and line charts, achieving 16% accuracy when tested with all 245 queries related to these chart types. Thus, the preliminary results indicate the significant potential of VizAbility for sighted users.

5 CONCLUSION & FUTURE WORK

This paper introduces a multimodal approach to enhance data visualization accessibility by combining keyboard navigation and speech interaction. The preliminary evaluation demonstrates the promising potential of our approach. In future work, we plan to improve performance by experimenting with advanced prompting strategies and conducting a comparative analysis of VizAbility against existing systems [6, 13], as well as other LLM services that currently support image input. Additionally, we aim to construct a larger benchmark dataset for a more comprehensive evaluation. Furthermore, we will conduct a user study involving blind participants to thoroughly assess the practical utility of VizAbility.

REFERENCES

[1] [n. d.]. Langchain. https://github.com/hwchase17/langchain. Accessed: August 21, 2023.

[2] Matt Blanco, Jonathan Zong, and Arvind Satyanarayan. 2022. Olli: An Extensible Visualization Library for Screen Reader Accessibility. In *IEEE VIS Posters*. http://vis.csail.mit.edu/pubs/olli

[3] Giovanni Fusco and Valerie S Morash. 2015. The tactile graphics helper: providing audio clarification for tactile graphics using machine vision. In *Proceedings of the 17th international ACM SIGACCESS conference on computers & accessibility*. 97–106.

[4] Cagatay Goncu and Kim Marriott. 2011. GraVVITAS: generic multi-touch presentation of accessible graphics. In *Human-Computer Interaction–INTERACT 2011: 13th IFIP TC 13 International Conference, Lisbon, Portugal, September 5-9, 2011, Proceedings, Part I 13*. Springer, 30–48.

[5] Shakila Cherise S Joyner, Amalia Riegelhuth, Kathleen Garrity, Yea-Seul Kim, and Nam Wook Kim. 2022. Visualization Accessibility in the Wild: Challenges Faced by Visualization Designers. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) *(CHI '22)*. Association for Computing Machinery, New York, NY, USA, Article 83, 19 pages. https://doi.org/10.1145/3491102.3517630

[6] Dae Hyun Kim, Enamul Hoque, and Maneesh Agrawala. 2020. Answering Questions about Charts and Generating Visual Explanations. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3313831.3376467

[7] 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*. 1–15.

[8] 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. https://doi.org/10.1111/cgf.14833 arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1111/cgf.14833

[9] N. W. Kim, S. C. Joyner, A. Riegelhuth, and Y. Kim. 2021. Accessible Visualization: Design Space, Opportunities, and Challenges. *Computer Graphics Forum* 40, 3 (2021), 173–188. https://doi.org/10.1111/cgf.14298 arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1111/cgf.14298

[10] OpenAI. [n. d.]. OpenAI API. https://openai.com/blog/openai-api. Accessed: August 21, 2023.

[11] Arvind Satyanarayan, Dominik Moritz, Kanit Wongsuphasawat, and Jeffrey Heer. 2016. Vega-lite: A grammar of interactive graphics. *IEEE transactions on visualization and computer graphics* 23, 1 (2016), 341–350.

[12] Ather Sharif, Sanjana Shivani Chintalapati, Jacob O. Wobbrock, and Katharina Reinecke. 2021. Understanding Screen-Reader Users' Experiences with Online Data Visualizations. In *Proceedings of the 23rd International ACM SIGACCESS Conference on Computers and Accessibility* (Virtual Event, USA) *(ASSETS '21)*. Association for Computing Machinery, New York, NY, USA, Article 14, 16 pages. https://doi.org/10.1145/3441852.3471202

[13] 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 *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) *(CHI '22)*. Association for Computing Machinery, New York, NY, USA, Article 478, 19 pages. https://doi.org/10.1145/3491102.3517431

[14] Alexa F. Siu, Danyang Fan, Gene S-H Kim, Hrishikesh V. Rao, Xavier Vazquez, Sile O'Modhrain, and Sean Follmer. 2021. COVID-19 Highlights the Issues Facing Blind and Visually Impaired People in Accessing Data on the Web. In *Proceedings of the 18th International Web for All Conference* (Ljubljana, Slovenia) *(W4A '21)*. Association for Computing Machinery, New York, NY, USA, Article 11, 15 pages. https://doi.org/10.1145/3430263.3452432

[15] Jason Wei, Xuezhi Wang, Dale Schuurmans, Maarten Bosma, Fei Xia, Ed Chi, Quoc V Le, Denny Zhou, et al. 2022. Chain-of-thought prompting elicits reasoning in large language models. *Advances in Neural Information Processing Systems* 35 (2022), 24824–24837.

[16] Markus Weninger, Gerald Ortner, Tobias Hahn, Olaf Drümmer, and Klaus Miesenberger. 2015. ASVG- Accessible Scalable Vector Graphics: intention trees to make charts more accessible and usable. *Journal of assistive technologies* 9, 4 (2015), 239–246.

[17] 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 (2022), 15–27. https://doi.org/10.1111/cgf.14519 arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1111/cgf.14519