In space, no one (but AI) can hear you scream

Mathis Brossier*
Linköping University, Sweden

Alexander Bock
Linköping University, Sweden

Konrad Schönborn
Linköping University, Sweden

Tobias Isenberg
Inria Saclay, France

Anders Ynnerman
Linköping University, Sweden

Lonni Besançon
Linköping University, Sweden

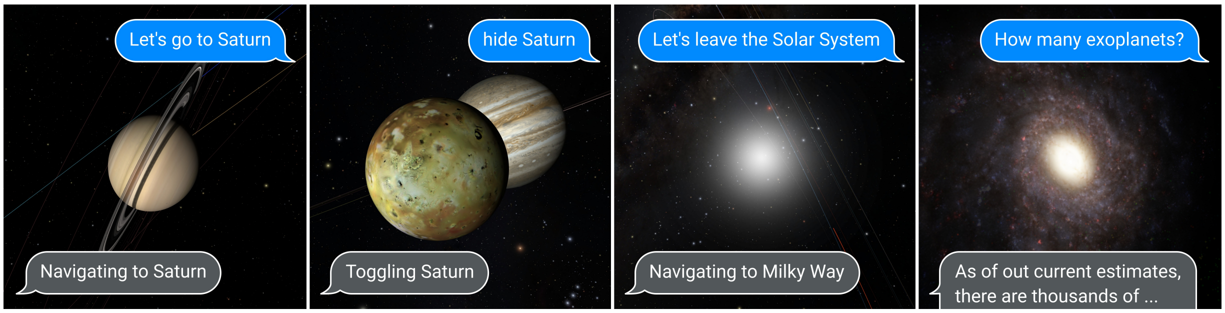

Figure 1: Examples of queries supported by our conversational agent with OpenSpace. Using the intuitive interface the user is able to navigate through the observable universe on a laptops or planetariums.

Abstract

We present our initial work on integrating a conversational agent using a large-language model (LLM) in OpenSpace, to provide conversational-based navigation for astrophysics visualization software. We focus on applications of visualization for education and outreach, where the versatility and intuitiveness of conversational agents can be leveraged to provide engaging and meaningful learning experiences. Visualization benefits from the development of LLMs by leveraging its capability to understand requests in natural language, allowing users to express complex tasks efficiently. Natural Language Interfaces can be combined with more traditional visualization interaction techniques, streamlining real-time interaction and facilitating free data exploration. We thus instructed a voice-controlled GPT-4o LLM to send commands to an OpenSpace instance, effectively providing the LLM with the the ability to steer the visualization software as a museum facilitator would for educational shows. We present our implementation and discuss future possibilities.

Index Terms: Visualization, Natural Language Interaction.

1 Introduction

Science museums increasingly rely on interactive visualization to enhance visitors' engagement and learning. Public spaces are crucial for enhancing learning and public interest [8] as well as promoting science to underserved communities [1]. Interaction facilitates a degree of freedom of exploration, which can tailor the visualization to specific interests of the public. Freedom of exploration alone, however, is not sufficient—it must be accompanied with proper guidance in the form of storytelling and points of interest [11].

Combining explorative and explanative visualization was coined as "*exploranation*" [4, 12]. A typical implementation of the concept is the astrophysical visualization software *OpenSpace* [3]. This platform allows viewers to roam freely in space and time and it showcases the state-of-the-art of our understanding of space in a unified environment that include, among other things, planets, stars, and satellites. The software is applied in planetariums during guided tours of the universe given to the broad public. While OpenSpace itself is highly interactive, the visitors/audience usually plays a passive role in these tours: the guide often follows a script, while a 'pilot' steers the visualization according to the story being narrated.

We aim to enhance the interactivity of guided visits, by empowering guides with unscripted interaction with the public. This process requires rapid and automatized ways to interact with visualization software—in our case OpenSpace, while presenting to an audience is still understudied [2]. We rely on Natural Language Interfaces (NLI) to facilitate the interaction. Indeed, with recent developments in Generative Artificial Intelligence and, in particular, Large Language Models (LLMs), NLIs can now facilitate many visualization tasks that would be tedious by traditional means, including tweaking the data and visuals on-the-fly (see Figure 1), tailoring visualization to specific tasks and users, and building narratives accompanied with a matching visualization to explain a topic.

2 Related work

Scholar have studied the use of natural language for visualization recommendation [6, 7], visual analytics [10], and interaction with visualization [9]. Jia *et al.*introduced VOICE [5], a conversational interface to a molecular visualization. Contrary to past efforts, this work focused on visualization and interaction with domain-specific scientific data for science communication. We build upon this work in the context of astrophysics, and integrate an NLI into pre-existing software with an established usage in science centers.

3 Scenario

As a first step towards integrating *conversation* in OpenSpace, we focus on a scenario in which a museum guide uses NLI to adapt during a planetarium show, based on the public's involvement. During the show the public may have questions, specific needs (e.g., a school field trip) or wish to react to recent events (e.g., a rocket launch). With the NLI, the guide may steer the visualization by themselves by sending voice commands, without interrupting their narration and without requiring coordination with the pilot.

*e-mail: mathis.brossier@liu.se

This approach allows us to safely test our implementation in real-world scenarios. It restricts direct access of the audience to the AI, but it allows visitors to play a more active role in the show nonetheless. The guide has an understanding of the AI abilities and limitations as well as of the inner workings of OpenSpace and can therefore, use the AI as a support tool only, and mitigating the risks of AI mistakes and hallucinations. This scenario combines the guide's proficiency in making science accessible with the AI's ability to react quickly to improvised scenarios.

4 Design & Implementation

OpenSpace provides a Python API that facilitates remote control of the software over the network. The LLM is initially prompted with a subset of the API's documentation, and instructions to call these functions to satisfy the user requests. We also provide The initial state of OpenSpace, including current location, time and available assets. We then use a separate process to (1) listen to the guide via a microphone controlled by a remote trigger, (2) convert voice to text with OpenAI's Whisper text-to-speech API, (3) prompt a GPT-4o assistant with the audio transcription, (4) execute the functions the LLM wants to call via the API, (5) send the function call results to the LLM.

Our system prompt instructs the LLM to execute any functions from the API it sees fit to best satisfy the user request. These functions include: Querying state information; traveling to a different target (e.g., planet, star, satellite); toggling the visibility of different assets; manipulating the view (zooming, panning, rotating); manipulating time (changing the date or the simulation speed); and answering general subject-matter questions. Based on the prompt, the LLM infers what action, or chain of actions would satisfy it. We provide the LLM with previous user prompts and responses as context to facilitate a conversational interaction. Communicating with voice can use two modes of operation—either the microphone continuously listens to the user and sends each sentence to the LLM, or the listening is triggered with a button. In our dome show context, a guide can easily use a presenter click er to trigger interaction only when needed.

5 Discussion

The system shows high proficiency in understanding direct requests to travel to a target ("*Go to Mars*") and explaining general concepts ("*Why is its surface red?*"). It also performs well in dialogue: understanding implicit questions based on a situational context ("*It's in the shadow*", while looking at a planet, is interpreted as "*rotate by 180°*"), seeking clarifications or expressing its inability to fulfill an request. For example, it shows poor understanding of space, time, and state changes. Questions such as "*What is the red dot?*", "*What is this island in the North-East?*" and "*What time is it?*", are challenging. Since the LLM is effectively blind, it relies on its internal representation of the visualization, which tends to drift away from reality as the discussion advances. This issue could be improved by providing the LLM with greater details on the current state, but also allowing the LLM to request more information to OpenSpace.

The LLM is currently able to execute explicit requests, but notbuild compelling personalized stories. In contrast, a guide knows how to maintain audience engagement, but cannot steer the visualization simultaneously. This is why we selected our scenario, as both parties (the AI—assistant and the human—expert) complement each other. In the future we will explore other applications of NLIs in science museums, and improve the robustness of the system for direct interaction with the NLI. This will require developing storytelling capabilities for the LLM to answer very indirect questions. For instance, a user may ask "*tell me the story about the solar system*". A proper response to such open-ended questions would have to be assertive and conveyed in multiple steps, each including a voice narration accompanied with a steering of the visualization, potentially even coupled with other generated visualizations such as custom charts (e.g., [7], as mentioned in previous work [5, 6]).

6 Conclusion & Future Work

We have presented a prototype combining generative AI and visualization for interactive educational and outreach about space. We hope that it can serve as a starting point to elucidate how such a conversational agent can be useful in this specific context through an in-situ study of our own dome shows at Norrköping Visualization Center C. This could serve as a basis to study the pedagogical merits of multimodal interaction with complex datasets by gathering insights from both experts and laypeople.

Supplemental Materials

Supplemental materials are available on OSF (doi: 10.17605/osf.io/jcgzt; CC BY 4.0 license): (1) source code and instructions for the LLM assistant and (2) a video showcasing its use in a planetarium.

Acknowledgments

This work is supported by the Knut and Alice Wallenberg Foundation (KAW 2019.0024), the Marcus and Amalia Wallenberg Foundation (MAW 2023.0130), and efforts from Visualization Center C.

References

[1] L. Archer, E. Dawson, A. Seakins, J. DeWitt, S. Godec, and C. Whitby. "I'm being a man here": Urban boys' performances of masculinity and engagement with science during a science museum visit. *J Learn Sci*, 25(3):438–485, 2016. doi: 10/gctkfn

[2] L. Besançon, A. Ynnerman, D. F. Keefe, L. Yu, and T. Isenberg. The State of the Art of Spatial Interfaces for 3D Visualization. *Computer Graphics Forum*, 40(1):293–326, Feb. 2021. doi: 10/gjbpxp

[3] A. Bock, E. Axelsson, J. Costa, G. Payne, M. Acinapura, V. Trakinski, C. Emmart, C. Silva, C. Hansen, and A. Ynnerman. OpenSpace: A system for astrographics. *IEEE Trans Vis Comput Graph*, 26(1):633–642, 2020. doi: 10/ggsrbm

[4] G. Höst, K. Palmerius, and K. Schönborn. Nano for the public: An exploranation perspective. *IEEE Comput Graph Appl*, 40(2):32–42, 2020. doi: 10/gjbjrf

[5] D. Jia, A. Irger, L. Besancon, O. Strnad, D. Luo, J. Bjorklund, A. Ynnerman, and I. Viola. VOICE: Visual oracle for interaction, conversation, and explanation. arXiv preprint 2304.04083, 2024. doi: 10/m4ft

[6] G. Li, X. Wang, G. Aodeng, S. Zheng, Y. Zhang, C. Ou, S. Wang, and C. H. Liu. Visualization generation with large language models: An evaluation. arXiv preprint 2401.11255, 2024. doi: 10/m4fv

[7] A. Narechania, A. Srinivasan, and J. Stasko. NL4dv: A toolkit for generating analytic specifications for data visualization from natural language queries. *IEEE Trans Vis Comput Graph*, 27(2):369–379, 2021. doi: 10/ghgthp

[8] S. Schwan, A. Grajal, and D. Lewalter. Understanding and engagement in places of science experience: Science museums, science centers, zoos, and aquariums. *Educ Psychol*, 49(2):70–85, 2014. doi: 10/gh2j5p

[9] L. Shen, E. Shen, Y. Luo, X. Yang, X. Hu, X. Zhang, Z. Tai, and J. Wang. Towards natural language interfaces for data visualization: A survey. *IEEE Trans Vis Comput Graph*, 29(6):3121–3144, 2023. doi: 10/gr2vzx

[10] A. Wu, Y. Wang, X. Shu, D. Moritz, W. Cui, H. Zhang, D. Zhang, and H. Qu. AI4vis: Survey on artificial intelligence approaches for data visualization. *IEEE Trans Vis Comput Graph*, 28(12):5049–5070, 2022. doi: 10/grxn3d

[11] A. Ynnerman, P. Ljung, and A. Bock. Reaching broad audiences from a science center or museum setting. In M. Chen, H. Hauser, P. Rheingans, and G. Scheuermann, eds., *Foundations of Data Visualization*, pp. 341–364. Springer, Cham. doi: 10/gtz4t6

[12] A. Ynnerman, J. Löwgren, and L. Tibell. Exploranation: A new science communication paradigm. *IEEE Comput Graph Appl*, 38(3):13–20, 2018. doi: 10/gd8wmr