EG UK Computer Graphics & Visual Computing (2024)
A. Slingsby and D. Hunter (Editors)

From LitVis to Reasoning in Data Visualisation: A Research Plan

M. Hutchinson, R. Jianu, A. Slingsby and P. Madhyastha

Department of Computer Science, City University of London

Abstract

In this paper, we present a research plan to capture and analyse the reasoning behind visualisation design choices using litvis notebooks. These notebooks provide a detailed narrative of the visualisation process, combining code, visual outputs, and natural language. Our approach aims to use these notebooks to create a dataset that reflects visualisation design in practice, focussing on the reasoning processes behind design choices. We will evaluate state-of-the-art multimodal models on this dataset and will carefully analyse the capacity of the models to generate reasoning traces. We will then build models that are capable of generating valid reasoning traces for visualisations. This research will result in reliable models for data visualisation that are grounded in realistic visualisation design processes.

CCS Concepts

• ***Human-centered computing*** → ***Visualization;*** • ***Computing methodologies*** → ***Natural language processing;***

1. Introduction

Natural Language (NL) is an essential part of visualisation. It is utilised within data visualisations as titles, annotations, and labels [SSC*22, SH22]; alongside visualisations as a means of communication, storytelling, and exposition [SH10, KM13, LS21]; and more recently as interaction modality within Visual Analytics (VA) systems to help users query data and generate visualisations [SSL*22].

Current VA systems incorporating NL as an input or output modality often focus on the final visualisation design, neglecting the underlying reasoning processes that lead to these designs. This is in conflict with previous work characterising the visualisation design process, which emphasises the *why* alongside the *how* [MMAM14] - the reasoning is as important as the implementation.

Research in Natural Language Processing (NLP) has explored reasoning processes [YZTW23], but the resulting datasets are often contrived, with tasks that do not accurately represent real-world applications. There are some multimodal datasets combining vision and language tasks that address these issues to a certain extent. However, many still suffer from multiple problems, including systematic shortcuts that can be exploited by models [DCTC21]. There is a need for more robust and realistic multimodal datasets that reflect real-world reasoning tasks.

Literate Visualisation (LV) [WKD18] is an approach for visualisation design and production that emphasises design exposition (DE), the articulation of the reasoning behind specific design choices. LV is inherently multimodal, combining visual outputs with NL expressions to capture the visualisation design process. Litvis is an open-source platform that supports LV. Litvis 'notebook' documents often employ a narrative schema — a structured framework that guides users in documenting their design process.

We will create a dataset that captures the reasoning process behind visualisation design choices using litvis notebooks. We aim to use this dataset to test the capabilities of current state-of-the-art multimodal models and evaluate their effectiveness in understanding and reasoning about the visualisation design process.

In this work we will ask the following **research questions**:

- How can litvis notebooks be used to capture and analyse the reasoning process behind visualisation design choices?
- Can litvis notebooks be used to develop a more realistic automated reasoning dataset?

2. Data Collection Methodology

We will collect data from student submissions, which will be anonymised and will consist of about 800 litvis notebooks produced by undergraduate and master's level students learning data visualisation.

In their submissions, students choose a dataset(s) to analyse and produce a complete visualisation project. These projects span a diverse range of topics, such as the London cycle hire scheme, the relationship between Twitter activity and weather, the Six Nations Championship, and the history of industrialisation in Britain.

The litvis notebooks are markdown documents, consisting of text, and code blocks written in elm. Visualisations are rendered using the elm-vegalite package [Woo18]. Each notebook will be analysed to extract useful information that is critical for understanding

the visualisation design process. This includes code, visualisation images, visualisation specifications, and natural language, and the relationships between each of these elements.

Students followed a narrative schema designed for this project guiding them to:

- ask **research questions**;
- design **visualisations**;
- describe the resulting **insights**;
- and articulate the **design justification** behind their choices.

This narrative schema will be used as the foundation of our dataset structure. The dataset will also include metadata, encompassing the mark and textual feedback. We can use this structure and relationships between each part of it to conduct an in-depth analysis of the decision-making process in visualisation. This will provide insights into how visualisation is conducted in practice. This structured approach will also facilitate question answering, which can be used to evaluate the effectiveness of AI models in reasoning about visualisation.

2.1. Evaluation

To ensure the quality of our dataset, we will evaluate each of the litvis notebooks. We will use the feedback provided as a measure of the quality of each submission. We will only include notebooks that contain sufficient design exposition and visualisations to ensure that we are able to capture the visualisation process comprehensively.

Additionally, as mentioned, the performance of state-of-the-art multimodal AI models will be evaluated using this dataset. For this evaluation, we will employ a range of automated Natural Language Generation (NLG) metrics such as BLEU, BLEURT, CIDEr, and SPICE [SMK20] to measure the quality of the generated responses. We will also explore the use of multimodal metrics, such as VIFIDEL [MWS19], to evaluate the coherence and relevance of the models' outputs in relation to the visualisations. These metrics will provide an assessment of the models' performance, and allow for comparison to other models and datasets.

3. Conclusion

By capturing and analysing the reasoning processes behind visualisation design in practice, we aim to enhance the understanding of the visualisation process. This project seeks to create a dataset that serves as a more realistic, application-oriented resource for NLP research, offering insights that could significantly improve the training and fine-tuning of models for application in the field of visualisation.

Beyond the creation of the dataset, this research will explore whether litvis notebooks are a useful resource for research at the interface of NLP and visualisation. We will evaluate the quality of these notebooks and assess the usability of the final dataset. This project will develop a cohesive methodology for transforming collections of litvis notebooks into a structured format for analysis. If successful, this methodology could be extended to other, more extensive notebook platforms such as Observable, potentially allowing us to leverage a vast amount of real-world, user-generated content.

References

[DCTC21] DANCETTE C., CADENE R., TENEY D., CORD M.: Beyond Question-Based Biases: Assessing Multimodal Shortcut Learning in Visual Question Answering, Sept. 2021. doi:10.48550/arXiv.2104.03149. 1

[KM13] KOSARA R., MACKINLAY J.: Storytelling: The Next Step for Visualization. *Computer 46*, 5 (Jan. 2013), 44–50. doi:10.1109/MC.2013.36. 1

[LS21] LUNDGARD A., SATYANARAYAN A.: Accessible Visualization via Natural Language Descriptions: A Four-Level Model of Semantic Content. *IEEE Transactions on Visualization and Computer Graphics 28*, 1 (Sept. 2021), 1073–1083. doi:10.1109/TVCG.2021.3114770. 1

[MMAM14] MCKENNA S., MAZUR D., AGUTTER J., MEYER M.: Design Activity Framework for Visualization Design. *IEEE Transactions on Visualization and Computer Graphics 20*, 12 (Nov. 2014), 2191–2200. doi:10.1109/TVCG.2014.2346331. 1

[MWS19] MADHYASTHA P., WANG J., SPECIA L.: VIFIDEL: Evaluating the Visual Fidelity of Image Descriptions. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics* (July 2019), Association for Computational Linguistics, pp. 6539–6550. doi:10.18653/v1/P19-1654. 2

[SH10] SEGEL E., HEER J.: Narrative Visualization: Telling Stories with Data. *IEEE Transactions on Visualization and Computer Graphics 16*, 6 (Oct. 2010), 1139–1148. doi:10.1109/TVCG.2010.179. 1

[SH22] STOKES C., HEARST M.: Why More Text is (Often) Better: Themes from Reader Preferences for Integration of Charts and Text, Sept. 2022. doi:10.48550/arXiv.2209.10789. 1

[SMK20] SAI A. B., MOHANKUMAR A. K., KHAPRA M. M.: A Survey of Evaluation Metrics Used for NLG Systems, Oct. 2020. doi:10.48550/arXiv.2008.12009. 2

[SSC*22] STOKES C., SETLUR V., COGLEY B., SATYANARAYAN A., HEARST M.: Striking a Balance: Reader Takeaways and Preferences when Integrating Text and Charts. *IEEE Transactions on Visualization and Computer Graphics* (Sept. 2022), 1–11. doi:10.1109/TVCG.2022.3209383. 1

[SSL*22] SHEN L., SHEN E., LUO Y., YANG X., HU X., ZHANG X., TAI Z., WANG J.: Towards Natural Language Interfaces for Data Visualization: A Survey. *IEEE Transactions on Visualization and Computer Graphics 29*, 6 (Feb. 2022), 3121–3144. doi:10.1109/TVCG.2022.3148007. 1

[WKD18] WOOD J., KACHKAEV A., DYKES J.: Design Exposition with Literate Visualization. *IEEE Transactions on Visualization and Computer Graphics 25*, 1 (Aug. 2018), 759–768. doi:10.1109/TVCG.2018.2864836. 1

[Woo18] WOOD J.: gicentre/elm-vegalite, Sept. 2018. URL: https://github.com/gicentre/elm-vegalite. 1

[YZTW23] YU F., ZHANG H., TIWARI P., WANG B.: Natural Language Reasoning, A Survey, May 2023. doi:10.48550/arXiv.2303.14725. 1