Evaluation and Improvement of Narrative Visualizations

Andrea Lezcano Airaldi[1]

[1]Department of Informatics – Faculty of Exact Sciences
National University of the Northeast (UNNE)

alezcano.exa.unne.edu.ar

Abstract. *In the field of information visualization, storytelling techniques help to communicate facts and enhance comprehension. The use of data storytelling best practices can inform the process of creating narrative visualizations and increase the quality of charts used in software applications by improving aspects such as memorability or engagement, for instance, supporting end users in the decision-making process. The main goal of this doctoral research is to develop a method for assessing and improving the quality of narrative visualizations in software products, like scatter plots, line, bar, or pie charts, among others. This has included a case study and a systematic mapping study.*

1. Introduction

Data visualization has become essential to understand large datasets and communicate findings. As a subfield of visualization research (Post et al., 2003), Information Visualization focuses on the visual representations of abstract data (Keim et al., 2008; Munzner & Maguire, 2014), to enhance understanding and to support and amplify cognition (Card et al., 1999). Visualizations are increasingly used to tell compelling stories in different domains. There are studies that review the use of decision support systems in fields such as agriculture (Gutiérrez et al., 2019) where visualization plays a key role in assisting end-users to interpret the data, or environmental sciences (Grainger et al., 2016), where it is necessary to interact with actors outside of the scientific community and therefore presents the challenge of effective communication to enable users to generate actionable understanding. In more structured contexts, researchers can use these stories to support discussion, decision making and process analysis (Willett et al., 2011).

Storytelling has long been used as an effective way of conveying information and knowledge. Stories aid memory and recall by embedding information into characters, settings, relationships, and events (Henry Riche et al., 2018); hence the importance of integrating data visualization into narrative stories. Many experts in chart design provide guiding principles to create quality visualizations (Bertin, 1983; Nussbaumer Knaflic, 2015; Tufte, 2001). By leveraging the best practices stablished in the literature we can reduce the cognitive workload associated with chart comprehension (Dowding et al., 2018; Gilger, 2006) and prompt positive decision-making (Nussbaumer Knaflic, 2015).

In a software application, the implementation of data storytelling best practices for the design of narrative visualizations can significantly impact comprehension, memorability, and engagement of charts, and assist end users in the decision-making process by addressing issues related to the size and complexity of the data (Dimara et al., 2021). The resulting visualizations should, however, be supported by an evaluation model that allows developers to assess their quality and thus assure effectiveness to a certain degree.

In recent years, evaluation has become a central issue in the field of visualization. There is a diverse set of qualitative and quantitative methods for evaluating different aspects of data-driven stories (Carpendale, 2008). Some of these include controlled experiments, usability tests and case studies (Plaisant, 2004) where participants perform benchmark tasks and researchers collect measures like completion time, error rate and accuracy. These methods, however, only focus on data “facts” and fail to consider other aspects such as decision-making support, or insight generation (Dimara et al., 2018). As Wall et. al. (2019) point out, it would be valuable to have tools that can be deployed rapidly and iteratively during the design process to evaluate visualizations prior to conducting experiments with users.

2. Research Goals

The main purpose of this doctoral research is to "provide a method for evaluating the quality and assisting in the improvement of narrative visualizations". This goal further breaks down into three specific goals, addressing both methodological and technological aspects:

SG1: Identify the best practices for the design of effective narrative visualizations. In general, a visualization is considered effective if it helps people extract accurate information (Card et al., 1999) without further complexity (Haroz & Whitney, 2012).

SG2: Build an evaluation model based on the patterns and best practices found in SG1 to assess the quality of narrative visualizations.

SG3: Develop a software tool to assist in the implementation of the model, and additionally provide improvement recommendations.

The working hypothesis states that: “the development of an evaluation model for narrative visualizations that incorporates best practices and materializes them in a semi-automated support environment, can increase the value of visualizations.” In this case, the value is directly linked to decision-making support, which has been identified as the end goal of data visualization (Munzner & Maguire, 2014; Ware, 2020).

The rest of this paper is structured as follows: Section 3 provides a summary of the background and related work, Section 4 describes the methodology and reports the results obtained so far, Section 5 presents the expected contribution of this doctoral research and Section 6 outlines future work.

3. Background and Related Works

This section presents an overview of prior research relevant to our work and contextualizes the aims of the thesis.

3.1. Information Visualization and Storytelling

Over the last two decades there has been a growing interest in visual storytelling and data visualization to communicate findings. Visualization dashboards are commonly used for decision-making but can be insufficient for communication purposes. In a recent study, (Sarikaya et al., 2019) found that people in Business Intelligence often put screenshots of dashboards into slide presentations, suggesting a need for more powerful storytelling features.

Riche et. al. (Riche et al., 2018) define "data-driven stories" as stories that are either based on or contain data, visualized to support one or more intended messages, usually including annotations (labels, pointers, text) or narration. Kosara and Mackinlay (Kosara & MacKinlay, 2013) argue that data-driven storytelling is a natural next step for data analysis and visualization and a pivotal component for effective data exploration.

Many authors have discussed emerging opportunities and challenges in storytelling in the field of visualization. Kosara and Mackinlay (Kosara & MacKinlay, 2013) give an overview of the topic, highlighting the importance of storytelling for data analysis and presentation, as visualizations are increasingly used for decision making. In (Rhyne et al., 2015), the authors take a narrower view of what data storytelling involves, to facilitate discussion around it, and present a comprehensive view of the storytelling process. Ojo and Heravi (Ojo & Heravi, 2017) examined 44 cases of award-winning data stories to identify storytelling practices and characterize the tools and techniques to create them.

Other research discuss the issues regarding visualization and data representation in different fields. In this context, (Gorodov & Gubarev, 2013) describes specific problems in Big Data visualization and defines a set of approaches to avoid them. They provide a useful classification of existing methods for data visualization in application to Big Data. In a similar way, (Gotz & Borland, 2016) reports on the emerging challenges facing the health industry with respect to medical information visualization, as well as some of the major opportunities. The authors suggest that interactive data visualization could become an essential tool for a data-driven healthcare system.

3.2. Evaluation of Visualizations

Given the limitations of formal laboratory methods such as controlled experiments, many researchers seek to go beyond this evaluation approach to determine the effectiveness of a visualization (Wall et al., 2019) and argue that traditional evaluation metrics might not be sufficient in motivating cognitively efficient visualizations (Qu & Hullman, 2016).

Tory and Möller (2005) discuss the use of traditional methods against alternative evaluation techniques in HCI, such as focus groups, field studies and expert reviews, particularly, heuristics evaluations. They argue that such reviews are valuable ways to assess visualizations. North et al. propose an insight-based evaluation (Saraiya et al., 2005) to assess how well a visualization supports people gaining insight from the visual representation.

Lam et al (2012) provide an overview of the different types of evaluation scenarios, categorized into those for understanding data analysis processes and those which evaluate visualizations themselves. They base their categorization on questions and goals, rather than existing methods, encouraging the community to consider the context before choosing an evaluation method.

Regarding evaluation criteria, Bertini, Tatu and Keim (Bertini et al., 2011) presented a systematic analysis of quality metrics to support the exploration of high-dimensional data sets and defined a quality metrics pipeline. More recently, in (Amini et al., 2018) the authors argue that data stories must address different challenges depending on the context and provide a non-exhaustive set of criteria and evaluation methods.

Based on the review outlined above, our evaluation model would be based on the visualization scenario by Lam et al (2012), as we will focus on visualizations as a final product, rather than the process to create them. We intend to develop a software tool to support and automate the evaluation process, providing practical design recommendations based on the guidelines and criteria found in the literature. Our research covers four main types of visualizations, namely: line charts, bar charts, scatter plots, and pie charts, in addition to choropleth maps, area charts, bubble charts and treemaps. We focus on these visualizations as they are the most frequently occurring types, according to (Battle et al., 2018; Lee et al., 2017)

4. Research methodology and obtained results

The development of this doctoral thesis will be conducted following the Design Science (DS) framework (Johannesson & Perjons, 2014), an approach for the creation of artifacts in the form of models, methods and systems to solve problems in a given context. Fig. 1 shows an overview of the high-level activities in the framework. Given that the main goal of this research is the construction of an artifact (the evaluation model discussed in Section 2), the Design Science framework was an appropriate fit for the thesis.

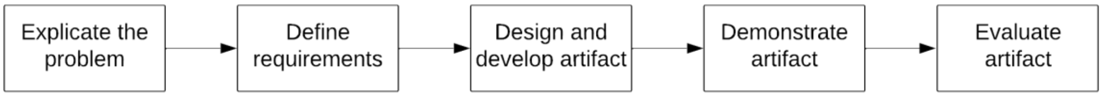

Fig. 1: Overview of the Design Science framework.

To address SG1 we are currently working on the first activity of the framework, that is, to explicate the problem. As a first task, we performed a case study following the guidelines by (Yin, 2017) and (Runeson & Höst, 2009) to gain a deeper understanding of the issue. The goal of the study was to determine the benefits of implementing data visualization best practices in the development of a software product, as well as the impact of not doing so. The system had high volatility in its requirements and short development cycles, prioritizing the delivery of functionality and real-time visualizations. This meant that narrative aspects were set aside and that only some best practices were followed by the development team. The results showed that the use of storytelling techniques in data visualization contributed to an improved decision-making process in terms of increasing information comprehension and memorability by the system stakeholders. In addition, we found that one of the reasons for not applying visualization best practices was the lack of knowledge regarding information visualization and storytelling, rather than the lack of time. In the absence of these skills, developers rely on the default settings of tools, causing stories to lose their potential or become difficult to understand. The full article is published at the Journal of Universal Computer Science and available online (Lezcano Airaldi et al., 2021)

As a second task, we are performing a systematic mapping study (SMS) following the guidelines by (Kitchenham & Charters, 2007), and (Petersen et al., 2015). For this purpose, we formulated the following research questions:

RQ1: *What are the existing definitions of "data storytelling"?* This question seeks to establish what data storytelling is and whether it has a formal and accepted definition for both academics and practitioners.

RQ2: *What are the data storytelling best practices reported in the literature for the design of visualizations and how are they applied?* The goal of this question is to summarize the guidelines and recommendations reported in the literature to create effective narrative visualizations and the diverse ways to apply them.

RQ3: *What are the criteria to evaluate narrative visualizations?* Depending on the goal of a visualization, different criteria is used to assess if that goal has been met (Amini et al., 2018). This question aims at identifying those criteria.

RQ4: *What are the current strategies to evaluate narrative visualizations?* This question identifies the methods by which the criteria defined in the previous question can be evaluated and their characterization: what types of charts they apply to, the metrics they use and the tools to support them.

Due to space limitations, the protocol for the SMS is detailed in Supplementary Materials[1].

The remaining activities in the DS framework will focus on SG2 and SG3. The definition of requirements will be done via a focus group to elicit requirements from stakeholders with different backgrounds and perspectives. For the design and development of the artifact, (Johannesson & Perjons, 2014) suggest drawing upon creative methods such as brainstorming sessions to generate ideas that could later be part of the design, in addition to software engineering techniques to develop it. To demonstrate the artifact, we plan on conducting an action research study to assess its feasibility. Finally, to evaluate the artifact, we will conduct two empirical studies: a controlled experiment that will allow us to achieve high internal validity (Johannesson & Perjons, 2014), and a case study to examine the artifact in a real world context.

5. Expected contribution

The main contribution of this research is an evaluation model implemented via a software tool to assess the quality of narrative visualizations. This model is intended to assist researchers and practitioners from various domains during the design phase of visualizations while also suggesting improvement recommendations based on the best practices and criteria found in the literature.

As an intermediate outcome, the contribution of the SMS is an extensive list of best practices and their corresponding applications that designers and developers can follow during the planning and design of visualizations or use them to compare to the ones they are currently adopting and identify improvement opportunities. Additionally, by understanding the criteria for effective visualizations, they can determine their goals

---

[1] https://bit.ly/cibse_appendix

more clearly (i.e.: to make visualizations more memorable, more comprehensible, or more engaging, for instance) and make informed design choices towards that direction.

6. Future work

We are conducting a systematic mapping study as part of the first activity in the Design Science framework towards SG1. We are currently working on reporting and discussing the results.

Once the review is completed, we will address the subsequent activities discussed in Section 4, detailing the tasks encompassed in each step. This will allow to iteratively build the different layers of the evaluation model proposed. We plan on publishing the resulting articles in journals and conferences relevant to the research field.

References

Amini, F., Brehmer, M., Bolduan, G., Elmer, C., & Wiederkehr, B. (2018). Evaluating Data-Driven Stories and Storytelling Tools *. In *Data-Driven Storytelling* (pp. 249–286). https://doi.org/10.1201/9781315281575-11

Battle, L., Duan, P., Miranda, Z., Mukusheva, D., Chang, R., & Stonebraker, M. (2018). Beagle: Automated Extraction and Interpretation of Visualizations from the Web. *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. https://doi.org/10.1145/3173574

Bertin, J. (1983). *Semiology of graphics*.

Bertini, E., Tatu, A., & Keim, D. (2011). Quality metrics in high-dimensional data visualization: An overview and systematization. In *IEEE Transactions on Visualization and Computer Graphics* (Vol. 17, Issue 12, pp. 2203–2212). https://doi.org/10.1109/TVCG.2011.229

Card, S. K., Mackinlay, J. D., & Shneiderman, B. (1999). *Readings in information visualization : using vision to think*. Morgan Kaufmann Publishers.

Carpendale, S. (2008). Evaluating information visualizations. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *4950 LNCS*, 19–45. https://doi.org/10.1007/978-3-540-70956-5_2

Dimara, E., Bezerianos, A., & Dragicevic, P. (2018). Conceptual and Methodological Issues in Evaluating Multidimensional Visualizations for Decision Support. *IEEE Transactions on Visualization and Computer Graphics*, *24*(1), 749–759. https://doi.org/10.1109/TVCG.2017.2745138

Dimara, E., Zhang, H., Tory, M., & Franconeri, S. (2021). The Unmet Data Visualization Needs of Decision Makers within Organizations. *IEEE Transactions on Visualization and Computer Graphics*. https://doi.org/10.1109/TVCG.2021.3074023

Dowding, D., Merrill, J. A., Onorato, N., Barrón, Y., Rosati, R. J., & Russell, D. (2018). The impact of home care nurses' numeracy and graph literacy on comprehension of visual display information: Implications for dashboard design. *Journal of the American Medical Informatics Association*, *25*(2), 175–182. https://doi.org/10.1093/jamia/ocx042

Gilger, M. (2006). Addressing information display weaknesses for situational awareness. *Proceedings - IEEE Military Communications Conference MILCOM*. https://doi.org/10.1109/MILCOM.2006.302129

Gorodov, E. Y. E., & Gubarev, V. V. E. (2013). Analytical review of data visualization methods in application to big data. *Journal of Electrical and Computer Engineering*, *2013*. https://doi.org/10.1155/2013/969458

Gotz, D., & Borland, D. (2016). Data-Driven Healthcare: Challenges and Opportunities for Interactive Visualization. *IEEE Computer Graphics and Applications*, *36*(3), 90–96. https://doi.org/10.1109/MCG.2016.59

Grainger, S., Mao, F., & Buytaert, W. (2016). Environmental data visualisation for non-scientific contexts: Literature review and design framework. *Environmental Modelling and Software*, *85*, 299–318. https://doi.org/10.1016/j.envsoft.2016.09.004

Gutiérrez, F., Htun, N. N., Schlenz, F., Kasimati, A., & Verbert, K. (2019). A review of visualisations in agricultural decision support systems: An HCI perspective. *Computers and Electronics in Agriculture*, *163*(May), 104844. https://doi.org/10.1016/j.compag.2019.05.053

Haroz, S., & Whitney, D. (2012). How capacity limits of attention influence information visualization effectiveness. *IEEE Transactions on Visualization and Computer Graphics*, *18*(12), 2402–2410. https://doi.org/10.1109/TVCG.2012.233

Henry Riche, N., Hurter, C., Diakopoulos, N., & Carpendale, S. (2018). Data-Driven Storytelling. In *Data-Driven Storytelling*. https://doi.org/10.1201/9781315281575

Johannesson, P., & Perjons, E. (2014). An introduction to design science. In *An Introduction to Design Science* (Vol. 9783319106328). Springer International Publishing. https://doi.org/10.1007/978-3-319-10632-8

Keim, D., Andrienko, G., Fekete, J. D., Görg, C., Kohlhammer, J., & Melançon, G. (2008). Visual analytics: Definition, process, and challenges. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *4950 LNCS*, 154–175. https://doi.org/10.1007/978-3-540-70956-5_7

Kitchenham, B., & Charters, S. M. (2007). Guidelines for performing systematic literature reviews in software engineering. In *Technical report, Ver. 2.3 EBSE Technical Report. EBSE* (Vol. 1).

Kosara, R., & MacKinlay, J. (2013). Storytelling: The next step for visualization. *Computer*, *46*(5), 44–50. https://doi.org/10.1109/MC.2013.36

Lam, H., Bertini, E., Isenberg, P., Plaisant, C., Carpendale Empirical, S., & Lam Enrico Bertini Petra Isenberg Catherine Plaisant Sheelagh Carpendale, H. (2012). Empirical Studies in Information Visualization: Seven Scenarios. *Institute of Electrical and Electronics Engineers*, *18*(9), 1520–1536. https://doi.org/10.1109/TVCG.2011.279ï

Lee, S., Kim, S. H., & Kwon, B. C. (2017). VLAT: Development of a Visualization Literacy Assessment Test. *IEEE Transactions on Visualization and Computer Graphics*, *23*(1), 551–560. https://doi.org/10.1109/TVCG.2016.2598920

Lezcano Airaldi, A., Diaz-Pace, J. A., & Irrazábal, E. (2021). Data-driven Storytelling to Support Decision Making in Crisis Settings: A Case Study. *JUCS - Journal of Universal Computer Science 27(10): 1046-1068*, *27*(10), 1046–1068. https://doi.org/10.3897/JUCS.66714

Munzner, T., & Maguire, E. (Graphic artist). (2014). Visualization analysis & design. In *Cs.Ubc.Ca*. A K Peters/CRC Press.

Nussbaumer Knaflic, C. (2015). *Storytelling with Data: A Data Visualization Guide for Business Professionals*. Wiley. https://www.amazon.com/-/es/Cole-Nussbaumer-Knaflic/dp/1119002257/ref=sr_1_1?crid=4WE4N3JTKA0L&dchild=1&keywords=storytelling+with+data&qid=1616074837&s=books&sprefix=storytelling%2Caps%2C315&sr=1

-1

Ojo, A., & Heravi, B. (2017). Patterns in Award Winning Data Storytelling. *Https://Doi.Org/10.1080/21670811.2017.1403291*, *6*(6), 693–718. https://doi.org/10.1080/21670811.2017.1403291

Petersen, K., Vakkalanka, S., & Kuzniarz, L. (2015). Guidelines for conducting systematic mapping studies in software engineering: An update. *Information and Software Technology*, *64*, 1–18. https://doi.org/10.1016/J.INFSOF.2015.03.007

Plaisant, C. (2004). The challenge of information visualization evaluation. *Proceedings of the Workshop on Advanced Visual Interfaces AVI*, 109–116. https://doi.org/10.1145/989863.989880

Post, F. H., Nielson, G. M., & Bonneau, G.-P. (2003). *Data visualization: the state of the art*. Kluwer Academic.

Qu, Z., & Hullman, J. (n.d.). *Evaluating Visualization Sets: Trade-offs Between Local Effectiveness and Global Consistency*. https://doi.org/10.1145/1235

Rhyne, T.-M., Lee, B., Riche, N. H., Isenberg, P., & Carpendale, S. (2015). *More Than Telling a Story: Transforming Data into Visually Shared Stories*. www.gapminder.org/

Riche, N. H., Hurter, C., Diakopoulos, N., & Carpendale, S. (2018). *Data-driven storytelling*. A K Peters/CRC Press.

Runeson, P., & Höst, M. (2009). Guidelines for conducting and reporting case study research in software engineering. *Empirical Software Engineering*, *14*(2), 131–164. https://doi.org/10.1007/s10664-008-9102-8

Saraiya, P., North, C., & Duca, K. (2005). An insight-based methodology for evaluating bioinformatics visualizations. *IEEE Transactions on Visualization and Computer Graphics*, *11*(4), 443–456. https://doi.org/10.1109/TVCG.2005.53

Sarikaya, A., Correll, M., Bartram, L., Tory, M., & Fisher, D. (2019). What do we talk about when we talk about dashboards? *IEEE Transactions on Visualization and Computer Graphics*, *25*(1), 682–692. https://doi.org/10.1109/TVCG.2018.2864903

Tory, M., & Möller, T. (2005). Evaluating visualizations: Do expert reviews work? *IEEE Computer Graphics and Applications*, *25*(5), 8–11. https://doi.org/10.1109/MCG.2005.102

Tufte, E. R. (2001). *The Visual Display of Quantitative Information* (2nd ed.). Graphics Press. https://www.edwardtufte.com/tufte/books_vdqi

Wall, E., Agnihotri, M., Matzen, L., Divis, K., Haass, M., Endert, A., & Stasko, J. (2019). A heuristic approach to value-driven evaluation of visualizations. *IEEE Transactions on Visualization and Computer Graphics*, *25*(1), 491–500. https://doi.org/10.1109/TVCG.2018.2865146

Ware, C. (2020). *Information visualization : perception for design* (4th ed.). Morgan Kaufmann.

Willett, W., Heer, J., Hellerstein, J. M., & Agrawala, M. (2011). CommentSpace: Structured support for collaborative visual analysis. *Conference on Human Factors in Computing Systems - Proceedings*, 3131–3140. https://doi.org/10.1145/1978942.1979407

Yin, R. K. (2017). Case Study Research and Design. In *Sage Publications* (6th ed.). SAGE Publications, Inc.