Visualization in the Wild

Christina Gillmann, *University of Leipzig, 04109, Leipzig, Germany*
Johanna Schmidt, *VRVis Zentrum für Virtual Reality und Visualisierung Forschungs-GmbH, 1220, Wien, Austria*
Daniel Wiegreffe, *University of Leipzig, 04109, Leipzig, Germany*

Visualization approaches have been successfully applied to a variety of application fields. Continuously published novel visualization approaches reflect the need for and success of visualization approaches. However, when looking at visualization techniques that designers apply in real-world applications targeted at a larger user group (e.g., medical experts, engineers, journalists, and public authorities), novel, published approaches hardly end up being used in real-world scenarios. The lack of readily available visualization software solutions and guidelines on applying these is unfortunate, especially since data science and data analytics are emerging fields, with many users needing data visualization. As shown in the different use-case scenarios outlined in this issue, implementing novel solutions within existing domain workflows is still a significant challenge for visualization researchers and domain experts. Challenges and guidelines for integrating interactive visualization applications in existing workflows are outlined in different real-world applications presented in this issue. The papers in this issue cover various domains and data types and comprise success stories of visualization approaches in different application domains.

The need for interactive visualization applications for understanding complex datasets and scenarios is accepted by experts from various different application domains.[7] This is also due to the fact that with the progress of digitization, enormous amounts of data are continuously collected, and experts somehow need to make use of this vast amount of information.[9] The ability to turn data into insights and actionable decisions plays an increasingly important key role in many different domains.[4]

However, the development of visualization solutions remains domain specific, with a lack for readily available and easy-to-use software tools.[8] Available visualization applications that can be used without programming skills are commercial and not freely available,[2] and are in many cases not suited for specific types of data (e.g., time series, geographic, and spatial data). The usage of Open Source visualization libraries is targeted toward users with programming skills,[1] which is often not the case for domain experts. Further, the application of visualization techniques lacks guidelines[5] on when to apply which technique to achieve a certain goal or fulfill a certain task.

Even the development of new visualization solutions for domain-specific problems is still a complex process. Models have been researched that describe this process, such as the nested model,[6] whose structure is very helpful in the development of new solutions. Only in the last years field studies for validation are increasingly conducted in a structured way.[3] Nevertheless, the validation by the adoption rate, which might be more interesting for the application areas, still plays a secondary role.[10]

VISUALIZATION IN THE WILD

We identified several factors that further enable the integration of interactive visualization in existing workflows, but also several factors that hinder the application of visualization.

Conducive Factors

Data understanding. In all cases, domain experts identified visualization as the primary technique to get insights into large amounts of data—even before the research project started. Over the

0272-1716 © 2023 IEEE
Digital Object Identifier 10.1109/MCG.2023.3317646
Date of current version 7 November 2023.

past years, the overall data and visualization literacy in the general public have increased due to the pervasiveness of visualizations during the pandemic and in everyday life.

Human in the loop: Especially when dealing with automated methods, domain experts agreed that visualization is needed to understand the output and relevance of automatically produced results. Human judgment is always required at the end, especially when decisions need to be made, and in this case, domain experts need to be able to compare with real-world data to judge and evaluate the automatically generated results.

Participation: Domain experts are willing to participate early in the visualization design process. Using low-fidelity prototypes and quick first results, using the data provided by the domain experts, increases the possibility for participation and brings the opportunity to define the final requirements better.

User experience: Users tend to use applications where they can see a clear benefit. If the benefit is large enough, they are willing to accept minor technical problems.

Hindering Factors

Data preparation: Data readiness is still seen as a significant challenge, and data preparation before being able to visualize it takes up a substantial amount of time in every project. Data preparation (often referred to as data wrangling) also comprises checking data quality and handling common data issues (e.g., missing data).

Trust: Although visualization literacy generally increased, domain experts are still reluctant to trust what they can see in a data visualization fully. The need for transparency, legibility, agency, and direct interactions with underlying data characterizes the agenda of the emerging field of data visualization. Increasing trust in data visualization can also mean, for example, to include additional statistical measures for interpretation.

Wealth of information: When getting to know a newly developed tool, users are easily overwhelmed by multiwindow applications. The wealth of information can be caused by a visualization design that tries to show as many different aspects of the data as possible. On the other hand, as datasets continuously grow and become more complex, the wealth of information can also be caused by complex use-case scenarios. In this case, users need time to interpret the information shown in the visualizations and put different parts of the displayed information into relation (e.g., real-world versus automatically generated data). It will be an interesting challenge in the future to support the aggregation of information and guide users to spot essential patterns, which could be, for example, supported by applying natural language models.

Lack of standard software tools: Domain experts, especially when not experienced in programming, still rely on readily available and easy-to-use applications like Excel for data analysis. The lack of visualization software being readily available hinders them from using more elaborate visualization techniques.

Visualization design versus analysis process: In some cases, the answers to questions asked by domain experts do not lead to a final visualization prototype at the end but rather show the importance of the design process, where domain experts are motivated to interact and further dive into their data, which is sometimes all that is needed to lead to a conclusion. In such cases, visualization only serves as a way for confirmation at the end.

ARTICLES IN THE SPECIAL ISSUE

We accepted four papers for this Special Issue through the formal *IEEE Computer Graphics and Applications* review process of the seven submissions. Lotteraner et al.[A1] present a design study together with hydrogeologists working on groundwater management. As a critical task, domain experts must continuously monitor and protect groundwater quality to ensure safe drinking water. In the design study, the authors explored the use of interactive tools to understand better the data collected at different sampling stations. The design study outlines the necessary steps to integrate interactive visualization into existing workflows for domain users who still need to gain experience with data analysis. The authors carefully outline the challenges during requirement engineering, the need for rapid prototyping, and the lack of standard visualization tools for connecting spatial and nonspatial data. Lo et al.[A2] researched on the user experience in augmented reality (AR) applications. The authors concentrated on an AR app that runs on smartphones and can be used by a novice audience to retrieve additional information during large-scale

APPENDIX: RELATED ARTICLES

A1. L. Lotteraner, T. Homann, and T. Möller, "The challenge of interdisciplinarity at the intersection of groundwater management and visualization research," *IEEE Comput. Graphics Appl.*, vol. 43, no. 6, pp. 50–63, Nov./Dec. 2023, doi: 10.1109/MCG.2023.3309090.

A2. W. H. Lo, H. Regenbrecht, and S. Zollmann, "Sports visualization in the wild: The impact of technical factors on user experience in augmented reality sports spectating," *IEEE Comput. Graphics Appl.*, vol. 43, no. 6, pp. 64–74, Nov./Dec. 2023, doi: 10.1109/MCG.2023.3308958.

A3. S. Nowak, B. A. Aseniero, L. Bartram, T. Grossman, G. Fitzmaurice, and J. Matejka, "Identifying visualization opportunities to help architects manage the complexity of building codes," *IEEE Comput. Graphics Appl.*, vol. 43, no. 6, pp. 75–86, Nov./Dec. 2023, doi: 10.1109/MCG.2023.3307971.

A4. L. Feierl, T. Möller, and P. Luidolt, "SunScreen: Visual fault detection for solar-thermal systems," *IEEE Comput. Graphics Appl.*, vol. 43, no. 6, pp. 87–100, Nov./Dec. 2023, doi: 10.1109/MCG.2023.3308962.

sports events. The authors relied on a controlled lab environment based on recordings of rugby games to run the study. The authors found out that some technical factors like jittering do not impact the user experience in AR environments. Other factors related to content, like registration accuracy (i.e., pointing to wrong players), very strongly affected the positive experience when using the application. The authors outline the challenges of integrating AR into real-world applications, including distance to objects and ergonomics. Nowak et al.[A3] explore the usage of interactive visualization for architects during the design process. So far, architects treated building codes as separate information during the design process. The only loose integration of building codes during the design process is unfortunate. Building codes are critical to architectural design as they define legal constraints and standards regarding safety, accessibility, sustainability, and others. The authors identified several stages within the workflow where interactive visualization could be helpful, for example, to maintain awareness of building codes during the design process. However, they also identified hindrances for applying visualization, which include trust and a general fear of being constrained in creativity in the design process. Feierl et al.[A4] explore the usage of interactive data visualization for fault detection in solar–thermal panels. Monitoring the performance of solar–thermal panels is crucial to ensure proper long-term operation. The data used for monitoring are very complex (i.e., time-dependent, multidimensional, and with different sampling intervals) and, thus, challenging to interpret, for which automatic fault detection algorithms are prone to raise false alarms. The authors outline a visual analytics framework development to support domain experts in fault detection based on time series data. During the evaluation, the authors identified common problems when applying new interfaces, like nonvisualization experts being easily overwhelmed with multiwindow applications and the interpretation of time series data and algorithm results simultaneously.

ACKNOWLEDGMENTS

The authors would like to thank Pak Chung Wong, the *IEEE Computer Graphics and Applications* Associate Editor-in-Chief of Special Issues, for his help throughout the process of creating this Special Issue. VRVis is funded by BMK, BMAW, Styria, SFG, Tyrol, and Vienna Business Agency in the scope of COMET—Competence Centers for Excellent Technologies under Grant 879730, which is managed by FFG. This work was supported in part by the Federal Ministry of Education and Research of Germany and in part by the Sächsische Staatsministerium für Wissenschaft Kultur und Tourismus in the program Center of Excellence for AI-research "Center for Scalable Data Analytics and Artificial Intelligence Dresden/Leipzig," project identification number: ScaDS.AI.

REFERENCES

1. P. Barlas, I. Lanning, and C. Heavey, "A survey of open source data science tools," *Int. J. Intell. Comput. Cybern.*, vol. 8, no. 3, pp. 232–261, 2015, doi: 10.1108/IJICC-07-2014-0031.
2. M. Behrisch et al., "Commercial visual analytics systems–Advances in the big data analytics field," *IEEE Trans. Vis. Comput. Graphics*, vol. 25, no. 10, pp. 3011–3031, Oct. 2019, doi: 10.1109/TVCG.2018.2859973.
3. S. Carpendale. *Evaluating Information Visualizations*. Berlin, Germany: Springer, 2008, pp. 19–45, doi: 10.1007/978-3-540-70956-5_2.
4. L. Cibulski, J. Schmidt, and W. Aigner, "Reflections on visualization research projects in the manufacturing industry," *IEEE Comput. Graphics Appl.*, vol. 42, no. 2, pp. 21–32, Mar./Apr. 2022, doi: 10.1109/MCG.2022.3156846.

5. E. Hehman and S. Y. Xie, "Doing better data visualization," *Adv. Methods Pract. Psychol. Sci.*, vol. 4, no. 4, 2021, Art. no. 25152459211045334, doi: 10.1177/25152459211045334.
6. T. Munzner, "A nested model for visualization design and validation," *IEEE Trans. Vis. Comput. Graphics*, vol. 15, no. 6, pp. 921–928, Nov./Dec. 2009, doi: 10.1109/TVCG.2009.111.
7. D. Rees and R. S. Laramee, "A survey of information visualization books," *Comput. Graphics Forum*, vol. 38, no. 1, pp. 610–646, 2019, doi: 10.1111/cgf.13595.
8. G. Reina et al., "The moving target of visualization software for an increasingly complex world," *Comput. Graphics*, vol. 87, pp. 12–29, 2020, doi: 10.1016/j.cag.2020.01.005.
9. G. Richer, A. Pister, M. Abdelaal, J.-D. Fekete, M. Sedlmair, and D. Weiskopf, "Scalability in visualization," *IEEE Trans. Visual. Comput. Graphics*, early access, Dec. 29, 2022, doi: 10.1109/TVCG.2022.3231230.
10. M. Tory and T. Möller, "Evaluating visualizations: Do expert reviews work?," *IEEE Comput. Graphics Appl.*, vol. 25, no. 5, pp. 8–11, Sep./Oct. 2005, doi: 10.1109/MCG.2005.102.

CHRISTINA GILLMANN is a researcher with the Signal and Image Processing Group, University of Leipzig, 04109, Leipzig, Germany, where she leads her own subgroup on UAVA in medical applications. Her research interests include UAVA, medical visualization, medical image analysis, and uncertainty analysis in general. Gillmann received her Ph.D. degree in computer science from the University of Kaiserslautern, Kaiserslautern, Germany. Contact her at gillmann@informatik.uni-leipzig.de.

JOHANNA SCHMIDT is the head of the Research Group and the coordinator of the respective research area *visual analytics* at the VRVis Zentrum für Virtual Reality und Visualisierung Forschungs-GmbH, 1220, Wein, Austria. Her main research interests include the visual analysis of time-dependent and heterogeneous data, and the integration of visualization in data science workflows in production and manufacturing environments. She is the corresponding author of this article. Contact her at johanna.schmidt@vrvis.at.

DANIEL WIEGREFFE is a senior researcher at the Image and Signal Processing Group, Leipzig University, 04109, Leipzig, Germany. His research interests include visual data analysis focusing on biological data, geographical data, and text data. Wiegreffe received his Ph.D. degree in computer science from Leipzig University. Contact him at daniel@informatik.uni-leipzig.de.