The Gap between Visualization Research and Visualization Software (VisGap) (2020)
C. Gillmann, M. Krone, G. Reina, T. Wischgoll (Editors)

Participatory Visualization Design as an Approach to Minimize the Gap between Research and Application

S. Jänicke[1], P. Kaur[2], P. Kuźmicki[1], and J. Schmidt[3]

[1]IMADA, University of Southern Denmark, Odense, Denmark
[2]Heinz Nixdorf Chair for Distributed Information Systems, Friedrich Schiller University Jena, Germany
[3]VRVis Zentrum für Virtual Reality und Visualisierung Forschungs-GmbH, Vienna, Austria

Abstract

Despite acceptance in our field, many sophisticated visualization projects suffer from failing acceptance by the targeted audience. Though the reasons for this circumstance might be manifold, we argue that they align with the typical pitfalls of software development. On the one hand, stakeholders are often not or only marginally integrated in the visualization design process, on the other hand, the goals we follow as visualization scholars do often not align with the goals of the stakeholders, reducing them to data deliverers. We provide case studies reporting on finished and ongoing projects following a participatory design approach. Those projects are initiated by the needs from users in digital humanities, biodiversity research, sports analysis and data science, and our results indicate that participatory visualization design leads to mutual benefits, reducing the gap between research and application in the targeted domain.

1. Introduction

The acceptance and sustainable integration of newly developed technical solutions by end users is a very complex and often long-running process. According to the **technical acceptance model** for software technologies [MG15], the usage of a new application primarily depends on two variables: (i) the *perceived usefulness* the new system has for the users, and (ii) the *perceived ease of use*. Only if users feel that a new system will give them easy access, while saving time and making work easier, will the new system continue to be used regularly. These two most important influencing factors for a sustainable application demonstrate that considering the views and expectations of the end users remains a major success factor of new technical systems.

As a consequence, the success of a new application can already be influenced while designing it by incorporating as much information on the intended use cases as possible. **Participatory design** (sometimes also referred to as co-operative design) is a design process where all stakeholders (e.g. partners, customers, end users) are actively involved in the design process [HP19]. This way participatory design ensures that the resulting technical solutions meet all needs, that the final systems are usable, and that it can be easily integrated into existing workflows of the end users. Participatory design is currently used in a variety of fields including software design [Kau10], architecture [Luc18], product design [GP02], and medicine [GHH*15]. The involvement of the end users in the process has always proven to be very rewarding and has lead to the creation of successful products [LJSO12].

Visualization researchers established predefined **development models for visualization applications** that designers follow to guide their work. The models try to capture the cyclic and flexible nature of the development process of visualization systems. Returning to previous stages and iterating through different stages is often necessary to achieve a valuable end result. Isenberg et al. [IZCC08] proposed a detailed evaluation stage at the beginning of the design process to assess the requirements of the visualization system to be designed. A similar approach has been proposed by Lloyd et al. [LD11] for geovisualization applications. Both approaches pertain to a task-based visualization development. The nested development model by Munzner [Mun09] proposes evaluation cycles in all stages of the development process. Such close collaboration with users is already close to the proposed model of participatory design. However, also in the nested development model the design of the visualization techniques to solve the use cases use is done by the visualization designer. The end users are then asked to evaluate the proposed visualizations, which might lead to a refinement of the system design.

We argue for a close collaboration of the end users in a **participatory design process** for visualization projects. In visualization, nowadays, interactions with end users and other stakeholders are usually intended for explaining the use cases and for evaluating the proposed visualization techniques. We propose a closer collaboration between all stakeholders in a visualization design process by considering the data workflows, technical environments, and sociological issues of the end users. This goes beyond task-based development models, as this design approach also considers

DOI: 10.2312/visgap.20201108

other factors that are important for increasing the acceptance of visualization systems. Of particular importance are feedback loops involving stakeholders in decisions at each development stage (see Figure 1). The contributions of this paper are twofold:

- **Case Studies.** The proposed participatory design approach for visualization has already been successfully applied in four exemplary case studies. The case studies describe successful implementations of visualization systems with close involvement of the end users.
- **Guidelines and Outlook.** We summarized the similarities of the case studies and the conclusions drawn therefrom into guidelines and suggestions for future research.

The four case studies are described in Section 2–5. A discussion of the benefits and problems of increased user involvement and an outlook on further research is given in Section 6.

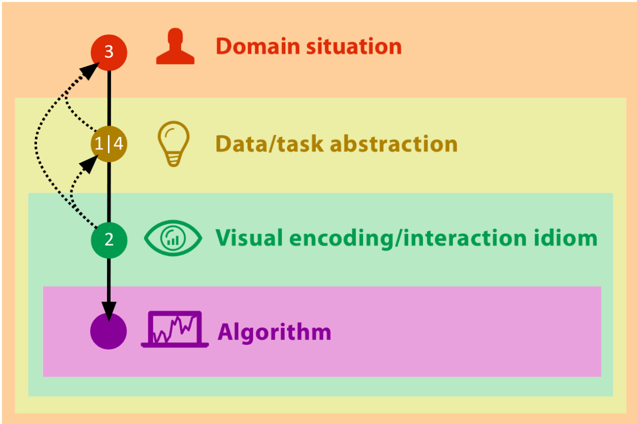

Figure 1: *By including feedback loops for all relevant decisions throughout a project from all levels of the model to previous levels, especially the domain situation, the nested model turns into a participatory visual design model (image adapted from [Mun09]). The numbers show entry points of the four digital humanities projects.*

2. Case Study 1: Digital Humanities

The digital humanities is an interdisciplinary research area that brings together scholars from different disciplines of the humanities and computer science. As opposed to qualitative humanities research, digital humanities focuses on empowering humanities scholars with computational means to carry out quantitative research questions on vast digitized cultural heritage collections [BDL*12]. Visualization is gaining more and more importance as a means to generate new perspectives on the data and to communicate occurring patterns, enabling to verify hypotheses and to generate new ones [JFCS17, WFS*19]. While interdisciplinary projects involving (digital) humanities and visualization scholars report on different collaboration setups and experiences [HEAB*17], there is a raising awareness of the benefits of methodological exchange in strong collaborations [BEAC*18].

Testimonial *by Stefan Jänicke*

I have been working in interdisciplinary digital humanities projects for twelve years. This testimonial continues my thoughts on the balancing act of valuable research on the intersection of digital humanities and visualization [Jän16] by reflecting on four of my projects in the light of Munzner's nested model [Mun09]. In particular, I recognized different layers of the model that served as entry points for the projects (see Figure 1). In addition, I report on measures to make the visualizations accessible to users and to increase their visibility.

1. **GeoTemCo** [JHS13] is a coordinated views framework that supports the comparative analysis of geospatial-temporal data. Its development in the *europeana-connect* [Eur20] was driven by an abstracted task, if supported by a visualization, perceived as potentially valuable for domain experts. Next to an existing data & task abstraction, a possible visual encoding was provided. Due to the misunderstood domain situation and changes in the technical specifications throughout the project years, the resulting *europeana4D* prototype was never used on a larger scale within the Europeana project. Further developments after the project turned *europeana4D* into *GeoTemCo* with an improved interface design to make to tool easier accessible to potential users. I made *GeoTemCo* publicly available at GitHub in 2012, and now it is used in the context of many scenarios, e.g., it appears as the GeoBrowser [GAU20] being one of the most important tools of the European digital humanities infrastructure project DARIAH.
2. **TRAViz** [JGF*15] is a text variant graph visualization that illustrates differences and similarities among different editions of a text on sentence- or paragraph-level developed during the eTRACES project [DHd20]. The interdisciplinary exchange with two philologists in the project team helped me understanding the domain situation. During my first attendance of the annual digital humanities conference in 2013, I received an even more precise picture of existing data abstractions [SC09] and interactive means demanded by domain experts [AVZ13]. However, standard tools such as CollateX [DM11] use standard graph visualization libraries that hardly communicate typical features of the variant graph data structure in visual form. TRAViz fills this gap. It is freely available and offers an easy-to-use interface leading to a widespread usage to align texts in different languages, e.g., [Ore17, BF19, Rus19]. Referring to Munzner's model, of particular importance was to not only the development of a visual encoding and interaction means easily comprehensible by domain experts, but also regular verifications if the requirements of data and task abstraction as well as the domain situation are met.
3. **MusikerProfiling** [JFS16] was my first project intentionally carried out on the basis of the nested model. It was initiated by a musicologist's research inquiry of discovering musicians having similar biographical characteristics. What followed was a participatory design process with a very frequent exchange, at times on a daily basis, on current developments and future steps. The underlying similarity measures were defined taking into account the domain expert's knowledge. The resulting visual analytics system supports the intended task, and it is accessible for musicologists worldwide since August 2015 through web-based interfaces for biographical databases of musicians. As of March 2020, the profiling interface was used by around

8,000 different users from over 70 countries performing more than 22,000 similarity and visual analyses concerning more than 6,000 different musicians. This interdisciplinary collaboration setup lead to a number of projects carried out in a similar fashion, e.g., [KKFJ19, KKFJ20].

4. **TagSpheres** [JS16] was rather a side product of an interdisciplinary project with different visions of the involved philologists, historians, NLP and visualization scholars. TagSpheres is a hierarchical tag cloud visualization that arranges co-occurrences dependent on their distances to a searched keyword spherically. Developing on basis of abstracting this task, it was finally only rarely used within the project as word order in the focused languages (Latin and Ancient Greek) was not crucial for many participating domain experts. In order to reach domain experts with more related research tasks, an online demonstrator of TagSpheres was prepared for and presented at the annual digital humanities conference in 2019 [JJG19]. This made the tool visible to the community and broadened the user base.

The nested model for visualization design has so far been an appropriate basis for my interdisciplinary projects. However, in order to create a visualization valuable for the targeted audience, stepping back to the domain situation always needs to be considered to ensure that a new visual design is intuitive and applicable, and that it accurately supports the intended user task. Moreover, as opposed to the original idea of the nested model should be laid out the way that domain experts are included in all decisions undertaken on different levels (see Figure 1) turning the nested model into a participatory visual design model. The major opportunity of this approach is that targeted users gradually build trust in the visualization. They benefit from being involved in the process of transforming and mapping raw data to a visual language, preempting that they regard the product as a black box, which would, especially in humanities applications, make a wide applicability impossible.

3. Case Study 2: Biodiversity Research

Biodiversity data is the data accumulated from the research done by biologists and ecologists on different taxa and levels, land use, and ecosystem processes. In order to answer the most relevant questions of biodiversity research, synthesis of data stemming from integration of data sets from different experiments or observation series is frequently needed. Collaborative projects thus tend to enforce centralized data management. This is true, e.g., for the *Biodiversity Exploratories* [FBG*10], a large-scale, long-term project funded by DFG. The Exploratories use the BExIS platform [LNB*12] for central data management. The instance of BExIS used within the Biodiversity Exploratiories (BE) serves as one of the primary sources for collecting requirements for this study. This data is highly complex, heterogeneous, and often not easy to understand. To search, explore, interpret, analyze, present, and reuse such data, a system is required to visualize these data sets effectively.

Testimonial ***by Pawandeep Kaur***

In the earlier years of my study to develop a visualization system for the BExIS data sets, I did many small scale requirement analysis studies with the biodiversity community. The aim for which was to know the visualization usage pattern of the community. Specifically, looking at the answers for: What visualizations are they aware of and visualization tools do they use? What problems do they face when they visualize their data, and what is their requirement from the visualization tool developed for the biodiversity data sets. Some of the answers and findings are summarized in the list below:

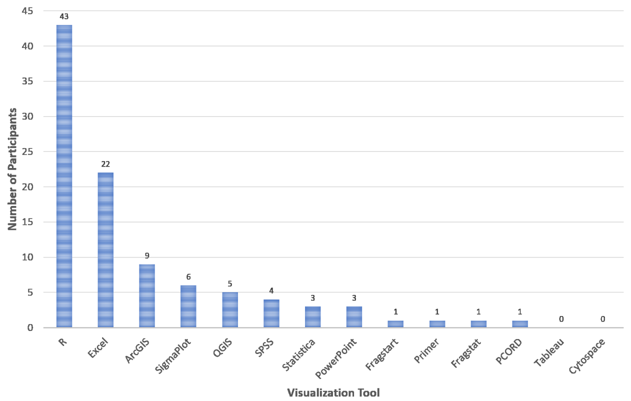

Figure 2: *Biodiversity use case. The bar chart shows a typical type of visualizations used by the biodiversity users. The data and the related chart are taken from my previous visualization requirement analysis survey [KKKR18].*

1. **Statistics and visualization tools**: A visualization tool for the biodiversity community must provide the functionality to do at least basic statistical analysis. In one of my previous surveys where the users were asked to select the visualization tools they use the most, the majority has voted for the software applications which provide statistical and data analytic functionalities (see Figure 2). The main task to visualize their data is to present the results of data analysis [KKKR18]. Therefore, regardless of the fact whether it needs programming or not, most of the community use tools with some statistical and analytical functionalities. Moreover, the most common visualizations are the one which shows the results of some statistical or data analysis, for example, ordination plots or biplots, scatter plot with regression a line, a bar chart with error bars, etc. As evident from my previous study [KKKR18], visualizations like parallel coordinates, treemaps, and other new visualization types are less accessible to them. The reason could be as they are comparatively new visualization techniques that are still not included in the traditional statistical software like SPSS [IBM20], etc. So the tools like Tableau [Tab20] that are primarily built for business analytics are not known to them.
2. **To know the goals of the biodiversity project and data well**: Biodiversity is a broad domain, and thus tools developed for one sub-domain of biodiversity may not be appropriate for others. For example, the goal of the 'species occurrence data' is to show the taxonomical, spatial, temporal change within the data set [GCA*13]. However, these goals may not hold for the 'species observation data' where the goal is to analyze the abundance and functional traits of the species that make up the ecosystem [JDLV18]. Furthermore, the goal of the raw biodiversity data is entirely different from that of biodiversity synthesis data. A raw data in biodiversity is a collection of ob-

servations for some species/biological or environmental entities (which was my case). Synthesis data sets are the one which is the integration of multiple data sets to answer the specific biodiversity questions. For example, to show the goal of 'network of species interactions, their connectivity, direction and intensity', data sets about species phylogenies, species interaction and species population at a particular location needs to be synthesized first. To understand the 'effect of land use and species functional traits', data about species phylogenies, species functional traits, land use effects etc. needs to be synthesized first. Thus understanding the specific goal of the biodiversity project and data is important to devise a visualization tools for them. It took me some time to know these differences and to integrate this knowledge into my work. I have realized that the visualization tool that I need to develop for the BExIS data sets is not primarily for presenting results of some analysis (as for synthesis data), but to visually explore the insights of the raw data.

These insights have reshaped my original research objectives. They have also provided me enough evidence that the visualization system for this community should understand the data domain as well as the context of the data. This leads me to the creation of the biodiversity domain knowledge-assisted visualization recommendation system [KK20]. This work-in-progress system is based on two important subsystems: (1) biodiversity text classifier that classifies the biodiversity text into different visualization and visual goals and (2) biodiversity context-aware variable selection algorithm that selects the important variables from the data set based on the context presented by the author in the related metadata.

4. Case Study 3: Football Data Analysis

The visualization of football game data has a long tradition, e.g., [RSB*10, PVF13, RKP18], but authors [SJB*16] also reported (at VIS 2017) on difficulties to deploy and establish sophisticated solutions in the targeted domain. This case study outlines a project involving computer and sport scientists as well as domain experts from the Danish Football League.

Testimonial *by Paweł Kuźmicki*

The Danish Football League (Superliga) contacted the University of Southern Denmark through one of our researchers at the health and sports department, who built many contacts to professional sports, now being the main contact person in ongoing collaborations involving the university. The project proposed by the Superliga required analysis, alignment, cleaning, improvement and visualization of game data acquired throughout the course of the previous football season. The contacted researcher's background is primarily sport and health science, so he included the department for mathematics and computer science in discussions with the Superliga. The following 2-stage project was assigned to the writer of this testimonial in the form of a master thesis project.

1. **Alignment ← Requiring expert knowledge**: The custom data sources consist of optical tracking data and event data. The optical tracking data describes all the X-, Y-and Z-positions of players, referees and a ball over a course of a football match in 25 frames per second. The event data lists all manually recorded actions that took place during the game (passes, shots, etc.). Those two data sets needed to be synchronized with each other as they are provided by two different external companies in different formats (CSV and XML). The synchronization required to map manually recorded events to the right frames in the tracking data set. The offered temporal information could only be used as a vague guess as the manual event information is incomplete, imprecise (delays of several seconds). Thus, typed events (more than 70 different types) were aligned to the tracking data taking positions and arrangements of players and the ball into account. In this phase of the project, whenever any football related question occurred, the sport and health researcher at the university essentially played a role of a football specialist. His expertise and knowledge of this field of study was sufficient to solve any arisen issues, i.e., to better get an understanding of how a specific event taking place on the pitch can be reflected in the tracking data.
2. **Enrichment → Addressing experts' needs**: When the project entered a more advanced stage, the exchange with the experts from the Superliga was more frequent. The league representatives guided the computer scientists in a development by expressing their desires regarding the data improvement and visualization, e.g., one of their propositions was a visualization of passes that send a ball into an opposite team's penalty box. The experts argue that a coach of a football team may benefit from this visualization by finding out which part of a pitch is a preferable way for the opposite team to enter the penalty box. This may in turn result in better selection of defenders in starting formation. Another wish expressed by the football experts were specific improvements of the data itself. The experts proposed a new set of events that would help them better understand a course of a football game. In close collaboration with the domain experts, the data source was enriched with reception (when a player receives the ball) and carry (the time a player possesses the ball) events, and an impact rate for passes.

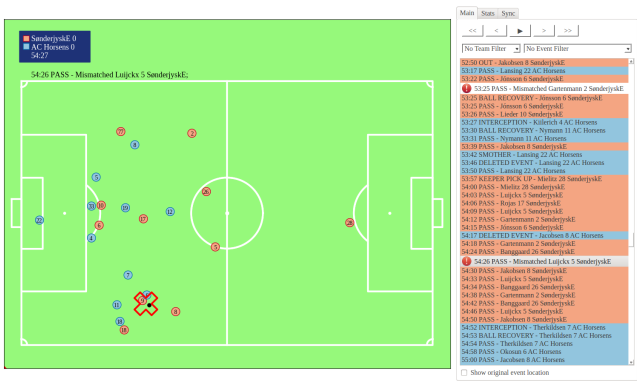

Figure 3: *2D animation of a football game highlighting mismatched alignments in both the pitch and the list.*

Figure 3 shows the resulting visual interface reporting on different events and mismatches in the event data, whose accuracy rate was of particular interest to the Superliga. Before the interdisciplinary exchange, machine learning techniques were planned to adopt in order to extract game situations. However, a remarkable reflection on the participatory design process was that the differing demands of the football experts reshaped the master thesis' topic.

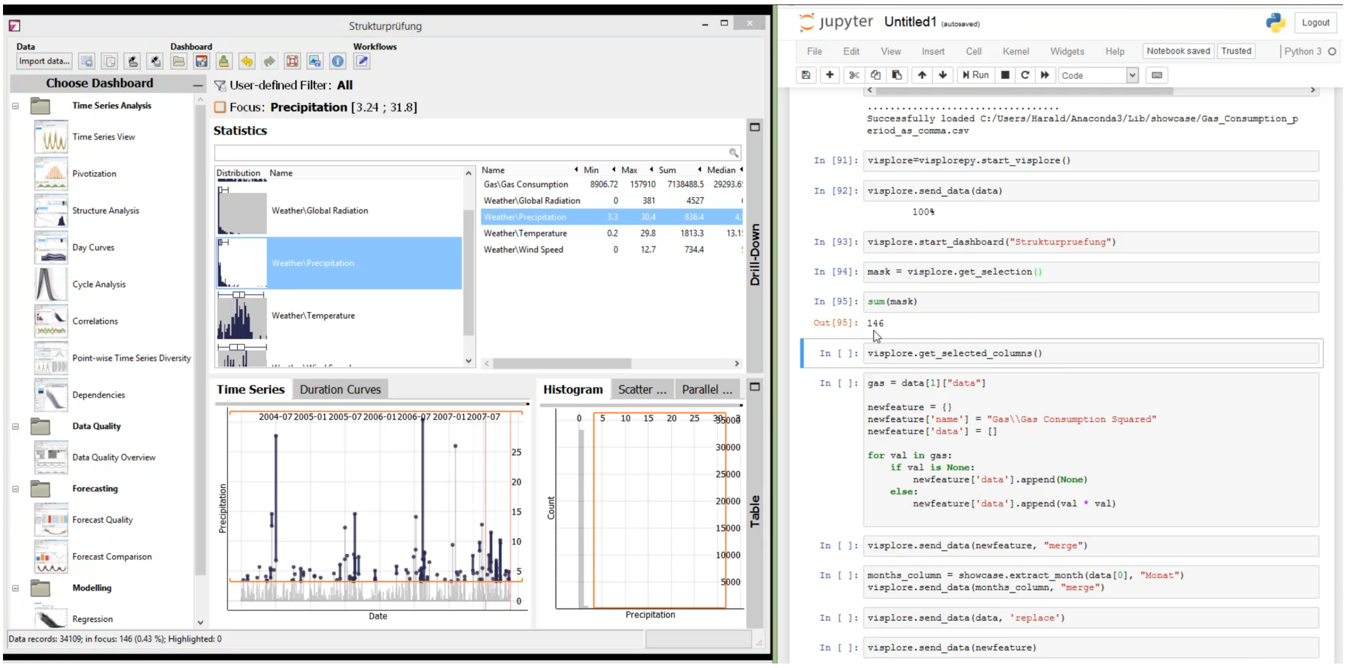

Figure 4: *Integrating visualization into scripting environments. The Visplore visual analytics system, which is a standalone application, provides interfaces to scripting environments (R, Python, Matlab) for allow a seamless integration in the data science workflows. Data can be sent to the Visplore system (left), and selections can be exported again into Python (right).*

5. Case Study 4: Data Science

Within the last years, data science has been established as an important emergent scientific field. Data science can be defined as a "*concept to unify statistics, data analysis, machine learning and their related methods*" in order to "*understand and analyze actual phenomena with data*" [Hay98]. As such, data science comprises the interdisciplinary integration of techniques from statistics, computer science, and information science [PGL*11]. Several studies were conducted within the last years to better understand the tasks and requirements of data scientists. The studies revealed that data science workflows are in many cases not targeted towards a clear goal and therefore require a lot of reconsideration and repetition [AZL*19] of tasks with different prerequisites.

Testimonial *by Johanna Schmidt*

The implementation of visual analytics tools was for many years driven by the idea of developing standalone applications that feature the whole pipeline of data import to visualization and reporting. This definitely supports users who are familiar with the data context, but are not familiar with using computational tools to analyze the data [KS12]. When working closely with data scientists in industry, I very soon realized that the situation is different. Since data scientists need to understand the data analysis process in detail, standalone applications, which are often perceived as black boxes, are not likely to be preferred.

Data scientists have to go through several steps on their way towards the goal of getting new insights from data. Their workflow can be summarized into five high-level categories [KPHH12], namely *Discover* (finding suitable data sets), *Wrangle* (bringing the data into a desired format), *Profile* (assessing the quality of the data and understanding its structure), *Model* (model building), and *Report* (reporting the findings). All steps in the workflow contain circular processes where data scientists have to rethink actions they made and restart analysis processes from scratch. This highly interactive and circular structure lead to data scientists using several different tools in different stages of their workflow. In fact, data scientists like to switch between different applications to always get the best solution for their current tasks [LBE19].

In my work I am therefore concerned with how to better integrate visual analytics in data science workflows. The visualization techniques currently applied in data science are quite limited, mostly only basic charts and plots are applied [Sch20]. This stands in a very strong contrast to the multitude of visualization techniques being developed [ML17], and to the possibilities for interactive exploration of the data. The system Visplore [PTMB09] developed at the VRVis is a standalone system specifically designed for the visual analytics of large time series data. To better integrate the system into data science workflows, we followed the following steps:

1. **Identifying needs.** We were told by several data scientists that it is important for them to fully understand the analysis methods they use, and that black box solutions are not of interest for them. One important point is, that they stressed that it is not possible for them to do without the algorithmic possibilities provided by Python, R, and other scripting languages. They also like to implement their own algorithms, something they can easily achieve in scripting languages. Data scientists got used to use

several different tools in their workflow, and they are therefore also used to combine tools through interfaces.

2. **Interface design.** A solution to better integrate scripting environments and Visplore was to provide a bidirectional channel between the programming environment and the Visplore functionalities. Data can be directly sent to Visplore via scripting, to visualize it, and to be able to explore details. Selections and data changes made by the users can be exported to the scripting environment again. For example, sending data to Visplore and visualizing it can reveal outliers in a time series. Users can then select the outliers in Visplore and remove them from the data set. The missing values can then be replaced by interpolated values, and the this way newly created data set can be exported to the scripting environment again. The interaction between the scripting languages and Visplore is illustrated in Figure 4. The interfaces work for R, Python and Matlab, since we identified these interfaces to be the most important ones for data scientists.
3. **Results.** Only close collaboration with the data scientists working with the data revealed their needs and requirements they have on visualization applications and lead to the implementation of the bidirectional interfaces. The direct input of the users was of great help when designing the scripting interfaces, and helped us to concentrate on the essential tools which are needed by the end users. The feedback from data scientists using the Visplore interfaces in practice was quite positive.

Interfaces between the programming environments used by data scientists and the provided visualization solutions [MPG*14] increase the application of visualization in their analysis workflows. This has also been noted by recent work on the gap between exploratory visualization tools being developed and the techniques used in data exploration [BE18]. I therefore think that including data scientists in our research designs and considering their workflows and environments will be necessary in the future to not loose the contact to these types of users.

6. Discussion & Conclusion

Our experiences from the use cases described in this paper lead to the conclusion that an early and close collaboration with all stakeholders in the design process leads to more successful visualization applications being developed. From our experience, we derive the following suggestions:

Understanding the domain. Being educated in computer science and visualization, we are typically unaware of research interests and current workflows in the targeted domain. On the one hand, we should be open to attain at least a basic knowledge, ideally, fascination, for the matters on the other side. In addition, we should try to understand how domain experts use state-of-the-art tools, if present, for their daily work. In primarily analogue domains as the humanities, one needs to observe traditional workflows in order to suggest computationally supported alternatives.

Closing visualization knowledge gaps. Domain scientists are typically unaware of the diversity of visualization techniques. Their visualization selection options are limited to what they have developed earlier or what they have seen in previously published works. They often do not know many options beyond static statistical graphs. This not only leads to a repetitive use of similar visualization types, it also hinders the use of modern sophisticated visualization. Regular workshops, participatory design studies and showcasing can help educating users from the targeted domain to inform them on current trends in visualization research.

Early prototyping. It can be beneficial to develop a basic functional prototype before gathering requirements from domain users. This way, they are better able to provide feedback when they can see something in real rather than abstract, and they are better able to map their existing workflows and reflect on conceptual gaps. Thus, they are more indulged and provide input on how to amend it to make it more suitable for their requirements. This early prototyping approach has been proven successful especially when targeting ecologists and inexperienced humanities scholars.

Opening the black box. Our studies show that many users from different target domains are reluctant to apply visualizations if they are not able to translate the processes how the data has been transformed and visually encoded. Firstly, a participatory design approach involves domain experts in all relevant decisions, and they are consistently evaluating current prototypes which leads to a better understanding and acceptance of the developed tool. In case rather complex methods such as neural networks are applied, efforts in explainable artificial intelligence (XAI) have to be undertaken to make sophisticated solutions accessible to target users.

Engage in the domain. Having a ready-to-use tool designed during an interdisciplinary project does not necessarily reach domain users beyond the project participants. It has been shown to be effective to present applicable solutions at conferences of the targeted domain. This not only potentially increases the user base of a tool, we can further perceive existing visualization gaps in the domain, and we can offer our expertise for related tasks leading to novel research directions in our field.

Open-source policy. Many domains such as the humanities are not equipped with funds to buy proprietary software supporting their research workflows. Therefore, offering our tools under an open-source license or in the form of an open-access, ready-to-use web-based application can lead to mutual benefits. This is typically a time-consuming task that we are not always willing to execute. Our experiences in carrying out participatory design processes however show that the end product is typically of a good quality as the code base benefits from several development cycles, and the data interface is tailored to standards established in the target domains.

While participatory design as an approach to interdisciplinary projects asks for target user involvement at all stages of the development process, the above listed aspects are meant to make the exchange of researchers from different domains easier and to increase the likeliness that all project participants are engaged in the collaborative setting. For future work, we would like to further explore the findings from the case studies towards formulating a new design process for visual analytics systems. We would like to evaluate the proposed participatory design process in comparison with other design processes which have been proposed so far. Although it has to be highlighted that such settings are time-consuming for us, our experiences show that projects carried out in a participatory design approach can even provoke future related research directions easier to tackle as contentual and methodological cross-domain barriers have already been overcome.

Acknowledgments

VRVis is funded by BMK, BMDW, Styria, SFG and Vienna Business Agency in the scope of COMET - Competence Centers for Excellent Technologies (854174) which is managed by FFG. Pawandeep Kaur is funded by the Deutsche Forschungsgemeinschaft (DFG) Priority Program 1374 "Infrastructure-Biodiversity-Exploratories" (KO 2209 / 12-2). Stefan Jänicke thanks Tamara Munzner for granting permission to include a modified version of the nested model into the paper.

References

[AVZ13] ANDREWS T. L., VAN ZUNDERT J. J.: An Interactive Interface for Text Variant Graph Models. *Digital Humanities 2013* (2013), 89–91. 2

[AZL*19] ALSPAUGH S., ZOKAEI N., LIU A., JIN C., HEARST M. A.: Futzing and Moseying: Interviews with Professional Data Analysts on Exploration Practices. *IEEE Transactions on Visualization and Computer Graphics 25*, 1 (2019), 22–31. doi:10.1109/TVCG.2018.2865040. 5

[BDL*12] BURDICK A., DRUCKER J., LUNENFELD P., PRESNER T., SCHNAPP J.: *Digital_Humanities*. MIT Press, 2012. 2

[BE18] BATCH A., ELMQVIST N.: The Interactive Visualization Gap in Initial Exploratory Data Analysis. *IEEE Transactions on Visualization and Computer Graphics 24*, 1 (2018), 278–287. doi:10.1109/TVCG.2017.2743990. 6

[BEAC*18] BRADLEY A. J., EL-ASSADY M., COLES K., ALEXANDER E., CHEN M., COLLINS C., JÄNICKE S., WRISLEY D. J.: Visualization and the Digital Humanities: Moving toward Stronger Collaborations. *IEEE Computer Graphics and Applications 38*, 6 (2018), 26–38. doi:10.1109/MCG.2018.2878900. 2

[BF19] BOSSE A., FANTA W.: *Textgenese in der digitalen Edition*. De Gruyter, 2019. 2

[DHd20] DHD: Digital Humanities im deutschsprachigen Raum - Projekt eTraces. https://dig-hum.de/forschung/projekt/etraces, 2020. [retrieved 2020-03-07]. 2

[DM11] DEKKER R. H., MIDDELL G.: Computer-Supported Collation with CollateX: Managing Textual Variance in an Environment with Varying Requirements. *Supporting Digital Humanities* (2011). 2

[Eur20] EUROPEANACONNECT: EuropeanaConnect as a Best Practice Network. https://www.europeanaconnect.eu/, 2020. [retrieved 2020-03-09]. 2

[FBG*10] FISCHER M., BOSSDORF O., GOCKEL S., HÄNSEL F., HEMP A., HESSENMÖLLER D., KORTE G., NIESCHULZE J., PFEIFFER S., PRATI D., RENNER S., SCHÖNING I., SCHUMACHER U., WELLS K., BUSCOT F., KALKO E. K., LINSENMAIR K. E., SCHULZE E.-D., WEISSER W. W.: Implementing large-scale and long-term functional biodiversity research: The Biodiversity Exploratories. *Basic and Applied Ecology 11*, 6 (2010), 473–485. doi:10.1016/j.baae.2010.07.009. 3

[GAU20] GEORG-AUGUST-UNIVERSITÄT: DARIAH-DE Geo-Browser. https://geobrowser.de.dariah.eu/, 2020. [retrieved 2020-02-29]. 2

[GCA*13] GAIJI S., CHAVAN V., ARIÑO A. H., OTEGUI J., HOBERN D., SOOD R., ROBLES E.: Content assessment of the primary biodiversity data published through GBIF network: status, challenges and potentials. *Biodiversity Informatics 8*, 2 (2013). doi:10.17161/bi.v8i2.4124. 3

[GHH*15] GORDON M., HENDERSON R., HOLMES J. H., WOLTERS M. K., BENNETT I. M.: Participatory design of ehealth solutions for women from vulnerable populations with perinatal depression. *Journal of the American Medical Informatics Association 23*, 1 (2015), 105–109. doi:10.1093/jamia/ocv109. 1

[GP02] GRUDIN J., PRUITT J.: Personas, Participatory Design and Product Development: An Infrastructure for Engagement. In *Proceedings of Participation and Design Conference* (Malmo, Sweden, June 23–25 2002), PDC '02, pp. 144–161. 1

[Hay98] HAYASHI C.: What is Data Science ? Fundamental Concepts and a Heuristic Example. In *Data Science, Classification, and Related Methods* (1998), Springer Japan, pp. 40–51. doi:10.1007/978-4-431-65950-1_3. 5

[HEAB*17] HINRICHS U., EL-ASSADY M., BRADLEY A. J., FORLINI S., COLLINS C.: Risk the Drift! Stretching Disciplinary Boundaries through Critical Collaborations between the Humanities and Visualization. In *Proceedings of the IEEE VIS Workshop on Visualization for the Digital Humanities* (Phoenix, Arizona, USA, Oct 2 2017). 2

[HP19] HARTSON R., PYLA P.: Chapter 19 - Background: Design. In *The UX Book (Second Edition)*. Morgan Kaufmann, Boston, MA, USA, 2019, pp. 397–401. doi:10.1016/B978-0-12-805342-3.00019-9. 1

[IBM20] IBM: SPSS-Software. https://www.ibm.com/analytics/at/de/technology/spss/, 2020. [retrieved 2020-03-09]. 3

[IZCC08] ISENBERG P., ZUK T., COLLINS C., CARPENDALE S.: Grounded Evaluation of Information Visualizations. In *Proceedings of the Workshop on BEyond time and errors: novel evaLuation methods for Information Visualization* (Florence, Italy, Apr 5 2008), BELIV '08, ACM. doi:10.1145/1377966.1377974. 1

[Jän16] JÄNICKE S.: Valuable Research for Visualization and Digital Humanities: A Balancing Act. In *Proceedings of the IEEE VIS Workshop on Visualization for the Digital Humanities* (Baltimore, Maryland, USA, Oct 24 2016). 2

[JDLV18] JAILLARD B., DELEPORTE P., LOREAU M., VIOLLE C.: A combinatorial analysis using observational data identifies species that govern ecosystem functioning. *PloS one 13*, 8 (2018). doi:10.1371/journal.pone.0203681. 3

[JFCS17] JÄNICKE S., FRANZINI G., CHEEMA M. F., SCHEUERMANN G.: Visual Text Analysis in Digital Humanities. *Computer Graphics Forum 36*, 6 (2017), 226–250. doi:10.1111/cgf.12873. 2

[JFS16] JÄNICKE S., FOCHT J., SCHEUERMANN G.: Interactive Visual Profiling of Musicians. *IEEE Transactions on Visualization and Computer Graphics 22*, 1 (2016), 200–209. doi:10.1109/TVCG.2015.2467620. 2

[JGF*15] JÄNICKE S., GESSNER A., FRANZINI G., TERRAS M., MAHONY S., SCHEUERMANN G.: TRAViz: A Visualization for Variant Graphs. *Digital Scholarship in the Humanities 30*, suppl 1 (2015), i83–i99. doi:10.1093/llc/fqv049. 2

[JHS13] JÄNICKE S., HEINE C., SCHEUERMANN G.: GeoTemCo: Comparative Visualization of Geospatial-Temporal Data with Clutter Removal Based on Dynamic Delaunay Triangulations. *Computer Vision, Imaging and Computer Graphics. Theory and Application* (2013), 160–175. doi:10.1007/978-3-642-38241-3_11. 2

[JJG19] JÄNICKE S., JOHN M., GESSNER A.: The Value of Tag Cloud Visualizations for Textual Analysis. In *Proceedings of the Digital Humanities* (2019). 3

[JS16] JÄNICKE S., SCHEUERMANN G.: On the Visualization of Hierarchical Relations and Tree Structures with TagSpheres. In *Proceedings of the 11th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (Revised Selected Papers)* (Rome, Italy, Feb 27–29 2016), VISIGRAPP '16, SciTePress, pp. 199–219. doi:10.1007/978-3-319-64870-5_10. 3

[Kau10] KAUTZ K.: Participatory Design Activities and Agile Software Development. In *Proceedings of the IFIP WG 8.2/8.6 International Working Conference - Human Benefit through the Diffusion of Information Systems Design Science Research* (Perth, Australia, Mar 30—Apr 1 2010), TDIT '10, Springer Berlin Heidelberg, pp. 303–316. doi:10.1007/978-3-642-12113-5_18. 1

[KK20] KAUR P., KIESEL D.: Combining Image and Caption Analysis for Classifying Charts in Biodiversity Texts. In *Proceedings of the 15th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications* (Valletta, Malta, Feb. 27–29 2020), VISIGRAPP '20, SciTePress, pp. 157–168. doi:10.5220/0008946701570168. 4

[KKFJ19] KHULUSI R., KUSNICK J., FOCHT J., JÄNICKE S.: An Interactive Chart of Biography. In *Proceedings of the 12th IEEE Pacific Visualization Symposium* (Bangkok, Thailand, Apr 23–26 2019), PacificVis '19, pp. 257–266. doi:10.1109/PacificVis.2019.00038. 3

[KKFJ20] KUSNICK J., KHULUSI R., FOCHT J., JÄNICKE S.: A Timeline Metaphor for Analyzing the Relationships between Musical Instruments and Musical Pieces. In *Proceedings of the 15th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications* (Valletta, Malta, Feb. 27–29 2020), VISIGRAPP '20, SciTePress, pp. 240–251. doi:10.5220/0008990502400251. 3

[KKKR18] KAUR P., KLAN F., KÖNIG-RIES B.: Issues and Suggestions for the Development of a Biodiversity Data Visualization Support Tool. In *Proceedings of the EG/VGTC Conference on Visualization (Short Papers)* (Brno, Czech Republic, June 4–8 2018), EuroVis '18, pp. 73–77. doi:10.2312/eurovisshort.20181081. 3

[KPHH12] KANDEL S., PAEPCKE A., HELLERSTEIN J. M., HEER J.: Enterprise Data Analysis and Visualization: An Interview Study. *IEEE Transactions on Visualization and Computer Graphics 18*, 12 (2012), 2917–2926. doi:10.1109/TVCG.2012.219. 5

[KS12] KANG Y., STASKO J.: Examining the Use of a Visual Analytics System for Sensemaking Tasks: Case Studies with Domain Experts. *IEEE Transactions on Visualization and Computer Graphics 18*, 12 (2012), 2869–2878. doi:10.1109/TVCG.2012.224. 5

[LBE19] LIU J., BOUKHELIFA N., EAGAN J. R.: Understanding the Role of Alternatives in Data Analysis Practices. *IEEE Transactions on Visualization and Computer Graphics (Early Access)* (2019). doi:10.1109/TVCG.2019.2934593. 5

[LD11] LLOYD D., DYKES J.: Human-Centered Approaches in Geovisualization Design: Investigating Multiple Methods Through a Long-Term Case Study. *IEEE Transactions on Visualization and Computer Graphics 17*, 12 (2011), 2498–2507. doi:10.1109/TVCG.2011.209. 1

[LJSO12] LINDSAY S., JACKSON D., SCHOFIELD G., OLIVIER P.: Engaging Older People Using Participatory Design. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Austin, Texas, USA, May 5–10 2012), CHI '12, ACM, pp. 1199—1208. doi:10.1145/2207676.2208570. 1

[LNB*12] LOTZ T., NIESCHULZE J., BENDIX J., DOBBERMANN M., KÖNIG-RIES B.: Diverse or uniform? — Intercomparison of two major German project databases for interdisciplinary collaborative functional biodiversity research. *Ecological Informatics 8* (2012), 10–19. doi:10.1016/j.ecoinf.2011.11.004. 3

[Luc18] LUCK R.: Participatory design in architectural practice: Changing practices in future making in uncertain times. *Design Studies 59* (2018), 139–157. doi:10.1016/j.destud.2018.10.003. 1

[MG15] MARANGUNIUNDEFINED N., GRANIUNDEFINED A.: Technology Acceptance Model: A Literature Review from 1986 to 2013. *Universal Access in the Information Society 14*, 1 (2015), 81–95. doi:10.1007/s10209-014-0348-1. 1

[ML17] MCNABB L., LARAMEE R. S.: Survey of Surveys (SoS) - Mapping The Landscape of Survey Papers in Information Visualization. *Computer Graphics Forum 36* (2017), 589–617. doi:10.1111/cgf.13212. 5

[MPG*14] MÜHLBACHER T., PIRINGER H., GRATZL S., SEDLMAIR M., STREIT M.: Opening the Black Box: Strategies for Increased User Involvement in Existing Algorithm Implementations. *IEEE Transactions on Visualization and Computer Graphics 20*, 12 (2014), 1643–1652. doi:10.1109/TVCG.2014.2346578. 6

[Mun09] MUNZNER T.: A Nested Model for Visualization Design and Validation. *IEEE Transactions on Visualization and Computer Graphics 15*, 6 (2009), 921–928. doi:10.1109/TVCG.2009.111. 1, 2

[Ore17] OREKHOV B. V.: The Word of Igor's Regiment. http://nevmenandr.net/cgi-bin/slovo.py?it=a1&it=a2&it=a4&it=a5, 2017. [retrieved 2020-03-12]. 2

[PGL*11] PARSONS M. A., GODØY Ø., LEDREW E., DE BRUIN T. F., DANIS B., TOMLINSON S., CARLSON D.: A conceptual framework for managing very diverse data for complex, interdisciplinary science. *Journal of Information Science 37*, 6 (2011), 555–569. doi:10.1177/0165551511412705. 5

[PTMB09] PIRINGER H., TOMINSKI C., MUIGG P., BERGER W.: A Multi-Threading Architecture to Support Interactive Visual Exploration. *IEEE Transactions on Visualization and Computer Graphics 15*, 6 (2009), 1113–1120. doi:10.1109/TVCG.2009.110. 5

[PVF13] PERIN C., VUILLEMOT R., FEKETE J.-D.: SoccerStories: A Kick-off for Visual Soccer Analysis. *IEEE Transactions on Visualization and Computer Graphics 19*, 12 (2013), 2506–2515. doi:10.1109/TVCG.2013.192. 4

[RKP18] RYOO M., KIM N., PARK K.: Visual Analysis of Soccer Players and a Team. *Multimedia Tools and Applications 77*, 12 (2018), 15603–15623. doi:10.1007/s11042-017-5137-4. 4

[RSB*10] RUSU A., STOICA D., BURNS E., HAMPLE B., MCGARRY K., RUSSELL R.: Dynamic Visualizations for Soccer Statistical Analysis. In *Proceedings of the 14th International Conference Information Visualisation* (London, UK, July 26–29 2010), IV '10, pp. 207–212. doi:10.1109/IV.2010.39. 4

[Rus19] RUSINEK S.: Seeing Echoes: Visualizing Text Reuse in Correspondence. https://densitydesign.github.io/como-sprints/2016%20cost%20action/seeing-echoes/, 2019. [retrieved 2020-03-12]. 2

[SC09] SCHMIDT D., COLOMB R.: A Data Structure for Representing Multi-version Texts Online. *International Journal of Human-Computer Studies 67*, 6 (2009), 497–514. doi:10.1016/j.ijhcs.2009.02.001. 2

[Sch20] SCHMIDT J.: Usage of Visualization Techniques in Data Science Workflows. In *Proceedings of the 15th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications* (Valletta, Malta, Feb. 27–29 2020), VISIGRAPP '20, SciTePress, pp. 309–316. doi:10.5220/0009181903090316. 5

[SJB*16] STEIN M., JANETZKO H., BREITKREUTZ T., SEEBACHER D., SCHRECK T., GROSSNIKLAUS M., COUZIN I. D., KEIM D. A.: Director's Cut: Analysis and Annotation of Soccer Matches. *IEEE Computer Graphics and Applications 36*, 5 (2016), 50–60. doi:10.1109/MCG.2016.102. 4

[Tab20] TABLEAU: Tableau. https://www.tableau.com, 2020. [retrieved 2020-02-12]. 3

[WFS*19] WINDHAGER F., FEDERICO P., SCHREDER G., GLINKA K., DÖRK M., MIKSCH S., MAYR E.: Visualization of Cultural Heritage Collection Data: State of the Art and Future Challenges. *IEEE Transactions on Visualization and Computer Graphics 25*, 6 (2019), 2311–2330. doi:10.1109/TVCG.2018.2830759. 2