A Research-Teaching Guide for Visual Data Analysis in Digital Humanities

Stefan Jänicke(✉)

IMADA, University of Southern Denmark, Odense, Denmark
stjaenicke@imada.sdu.dk

Abstract. The use of visualization to underpin distant reading arguments on cultural heritage data has established in the digital humanities domain. Novel strategies to represent data visually typically arise from interdisciplinary projects involving humanities and visualization scholars. However, the quality of outcomes might be inhibited as typical challenges of interdisciplinary research arise, and, at the same time, problem solving strategies are missing. I taught a course on visual data analysis in the digital humanities to let students with diverse study backgrounds experience those challenges in their early academic careers. This paper illustrates the research-teaching components of my course. This includes the contents of the theoretical training with active learning tasks, aspects of the practical training and considerations for teachers aiming to compose a related course.

Keywords: Research-teaching nexus · Visual data analysis · Digital humanities

1 Introduction

The importance of visualizations in digital humanities applications has steadily increased in recent years [22]. However, it has been shown that many visual representations are produced from standard tools that are not necessarily tailored for the given research task [24]. On the other hand, interdisciplinary projects can suffer from misunderstandings leading to visual designs that are hardly applicable by domain experts. My intention to offer a visualization course that attracts both computer science and humanities students was to equip them with the necessary knowledge and terminology in order to be able to "speak the same language" in potential future projects on the intersection of visualization and digital humanities. Therefore, students collaboratively conducted visual design projects in which they applied learned theoretical contents and faced typical challenges of interdisciplinary research.

This paper is an extended version of an IVAPP conference paper [32]. In addition, it contains aspects of a report on teaching journalism students [25]. While these publications rather report on the experiences gained in three years of teaching the course with a detailed discussion of conducted research projects, this paper contributes the following aspects:

– *Research-teaching Nexus:* The paper discusses different components of the course, which mainly focuses on conveying the research processes that lead to knowledge construction, referring to the research-teaching nexus by Healey [17].

© Springer Nature Switzerland AG 2022
K. Bouatouch et al. (Eds.): VISIGRAPP 2020, CCIS 1474, pp. 205–222, 2022.
https://doi.org/10.1007/978-3-030-94893-1_9

– *Theoretical Training:* A structured overview of the taught theoretical material is provided, and active learning tasks for individual classes are proposed.
– *Practical Training:* This paper puts a stronger focus on the roles of the teacher in interdisciplinary student projects, and projects that contributed to publications are discussed in detail.
– *Considerations:* Dependent on the cohort of participating students, different aspects that influence structuring course contents are discussed.

2 Terminology

In analogy to my first class in which I present the big picture of the course, this section is dedicated to discussing the terms that compose the title of the paper.

2.1 Digital Humanities

The digital humanities is an interdisciplinary field on the intersection of the humanities and computer science aiming to systematically apply computer-based methods to digital cultural heritage resources. Including a large variety of traditionally grown domains like literary studies, political sciences or computer science, interdisciplinarily oriented fields like computational linguistics, digital history or applied informatics also fall under the umbrella of the digital humanities. This diversity of disciplines makes defining *the digital humanities* a hard task. Heyer and Isenberg [18] describe the terms computational and digital humanities as different strategies to work with digital material. While the digital humanities would focus on creating digital repositories for cultural heritage collections, computational humanities would provide computational methods to analyze and enrich those repositories. The term *eHumanities* stands somewhere in between. Looking at the history of the domain, I use the traditional *digital humanities* term to describe all activities related to gathering and processing digitized material.

A definition for the domain, not aiming to generate a strict boundary, can be rather derived from the methodologies that lead to knowledge construction in the associated areas of research [52]. Whereas the natural sciences are concerned with the systematic study of nature and the discovery of regularities with which (natural) phenomena can be explained, the humanities consider all those phenomena that result from human action, evaluating them in their significance for the individual and society, and thus serving to create human consciousness. Tying those two methodologies together, the *digital humanities* could be defined as an area that generates a quantitative, computationally driven perspective on cultural heritage data, aiming to derive regularities of human action that contribute to creating human consciousness.

This turn from a qualitative to a quantitative analysis of information is depicted in Fig. 1 by the example of comparative literature analysis. The traditional approach includes a close reading of few related works followed by a detailed analysis of the writers intentions. Digital humanities approaches, on the other side, feature distant readings of text. By disrupting the linearity of text and counting occurrences of words, rather universal insights can be gained. In the given example, a literary scholar performed a keyword-based searches on a large text corpus of Latin texts and compared the aggregated contexts in which three synonymous keywords appeared using a comparative

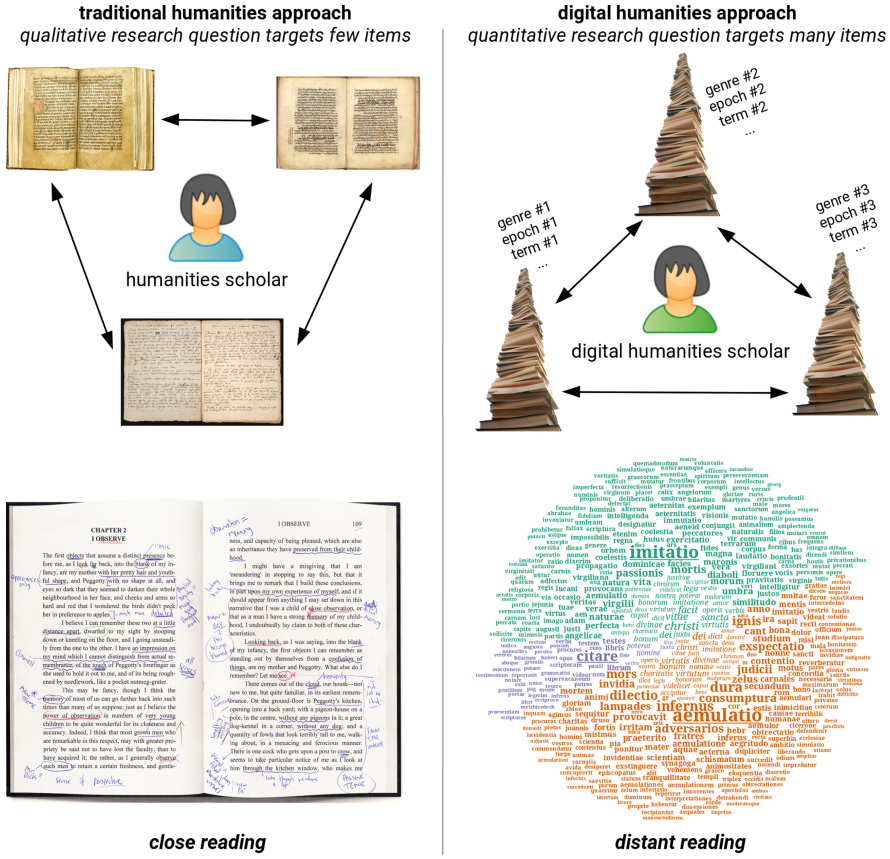

Fig. 1. Traditional and digital humanities approaches for comparative literature analysis. A traditional workflow would be based on close readings of few literary works (close reading image provided by Kehoe [36]), while digital humanities workflows typically involve quantitative distant readings [26].

tag cloud visualization [26]. By comparing shared and non-shared co-occurrences, the scholar was able to clarify and discriminate the meaning of words related to *imitate*. It has to be said that an analysis is typically not done only referring to a distant reading. Visualizations usually aid to generate different perspectives on information and access to close readings is granted interactively.

2.2 Visual Data Analysis

Recent survey papers document the increasing importance of visualizations for digital humanities research [22, 62]. Visual interfaces have become indispensable components to generate bridges between a computationally gained result and expert users. Tominski and Schumann describe interactive visual data analysis as a "synthesis of

visualization, interaction, and automatic computation to facilitate insight generation and knowledge crystallization from large and complex data" [58]. This synthesis is depicted in the visual data analysis process [22], which illustrates how insight is gained in digital humanities projects (see Fig. 2). A motivating data analysis task influences all components of the process. Data transformation, depending likewise on the underlying research task and the raw data format, deliver the input for visual interfaces used for visual analysis purposes.

The visual data analysis process in the context of digital humanities applications typically follows Shneiderman's Information Seeking Mantra *Overview first, zoom and filter, then details-on-demand.* Distant reading visualizations provide overviews, and means of zooming and filtering provide access to detailed close reading visualizations. The red arrows in Fig. 2 illustrate the iterative nature of visual data analysis in the context of digital humanities applications. Not seldomly, the process includes numerous cycles of adjustment or even redesign. There are two major reasons that induce modifications to an existing solutions. On the one hand, new perspectives on the data might lead to new ideas or a direction change in a project [20]. On the other hand, miscommunications typical for interdisciplinary projects lead to a visual design incapable of constructing knowledge concerning the data analysis task. This is especially crucial for digital humanities collaborations, in which research backgrounds are diverse. My motivation for teaching the course on visual data analysis in digital humanities is to bring together students from computer science and the humanities, so that they experience the challenges of interdisciplinary work early in their academic careers.

2.3 Research-Teaching Nexus

Given the clear focus on a targeted research domain for applying visualization and developing new techniques to visually represent information, the curriculum of my course can be described on the basis of the research-teaching nexus introduced by Healey [17]. As can be seen in Fig. 3, research-teaching can take different forms dependent on the targeted learning outcome and how students engage with the research materials. If a course is *research-led*, research contents and the important results of a domain are central to discussion. This is the rather traditional form of teaching that clearly separates the roles of the active teacher talking to a rather passive audience. Lecturing can also be *research-oriented*, which, instead of research findings, aims to convey research methodologies and best-practices that lead to knowledge construction concerning the taught subject. In *research-tutored* and *research-based* forms of teaching, students more actively participate and engage with research materials. While a *research-tutored* course involves activities such as reading and writing papers, in other words, a critical discussion of research contents, students practically experience research processes themselves when focusing on *research-based* activities.

The major intention of my course is to enable students to investigate research questions on the intersection of digital humanities delivering the research inquiry and visualization tailoring appropriate solutions to support related analysis tasks. Thus, the course focuses primarily on conveying how visual design models can be used to approach a given research task. Related examples from the literature underpin the value of established workflows for knowledge construction in this interdisciplinary field. Therefore, the theoretical training in the first half of the course is mainly *research-oriented*

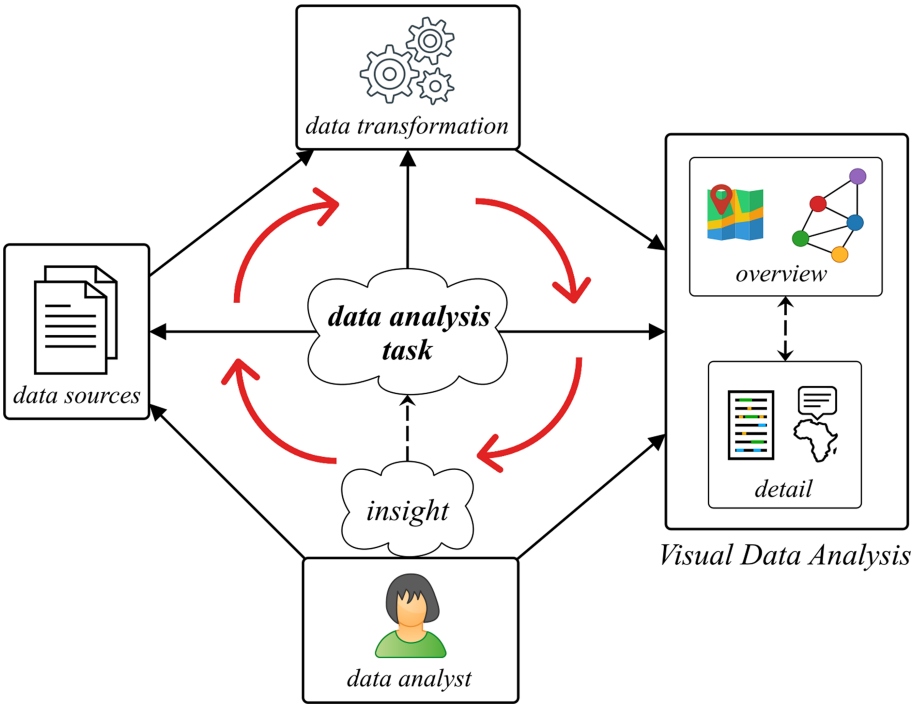

Fig. 2. Visual data analysis process in digital humanities based on [22].

as it teaches on how visualization can be best applied to existing problems to generate knowledge concerning cultural heritage data. The *research-led* part refers to the required basics of information visualization. In order to activate students during the theoretical training in a *research-based* fashion, students can replicate visual data analysis tasks on a small scale. The practical training conducted in the second half of the course is fully *research-based* as students experience real world research scenarios and develop means of visual analysis to investigate the given problem by following taught visual design approaches and workflows. The *research-tutored* part of the practical training refers to publishing promising results of course projects after the course.

3 Theoretical Training

According to the research-teaching nexus [17], the major focus of the theoretical training is to make students understand the processes that support knowledge construction in digital humanities with the aid of visual interfaces. Therefore, the theoretical training is mainly research-oriented. Research-led fragments of the course relate to the discussion of research findings on which visualization scholarship is based on, and research-based *active learning* tasks throughout the theoretical training support the understanding of the theoretical contents and prepare the students for their practical training conducted in the second half of the course.

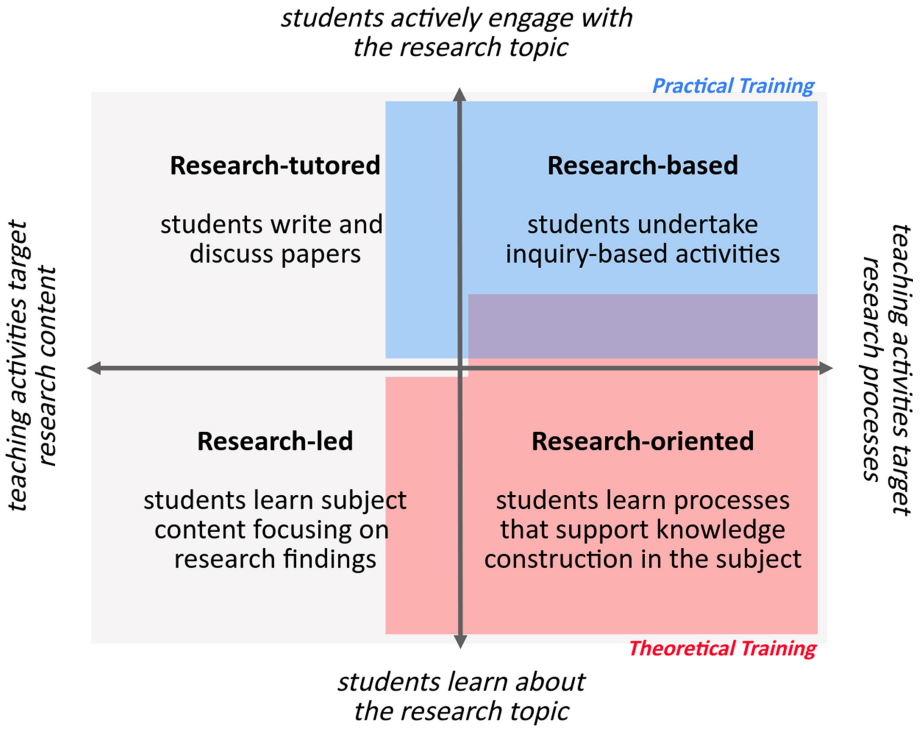

Fig. 3. Curriculum design and the research-teaching nexus based on Healey [17].

The schema of the research-oriented, theoretical training part of the course is depicted in Fig. 4. The syllabus adopts well-established strategies of teaching information visualization [37,51], however, with the digital humanities, focusing on a particular application domain. That means, the course The theoretical training focuses primarily on conveying the processes that lead to knowledge construction in digital humanities using visualization. Therefore, I structured the lecture contents in three blocks.

3.1 Block 1: The Big Picture

The course title included (in my case) the terms *Information Visualization* and *Digital Humanities*. I reserved the first class to give a broad overview of the individual terms and how they play together in the context of applying visualization in digital humanities. The discussion of sample projects that implement the visual data analysis process helped to convey the importance of the taught topic. We discussed four digital humanities projects, all based on raw textual sources, but each with a different research target in mind:

- Geospatial-temporal analysis of places mentioned in Herodotus' Histories with maps and timelines [4]
- Comparative analysis of similarities and differences among text editions with heat maps and variant graphs [29]

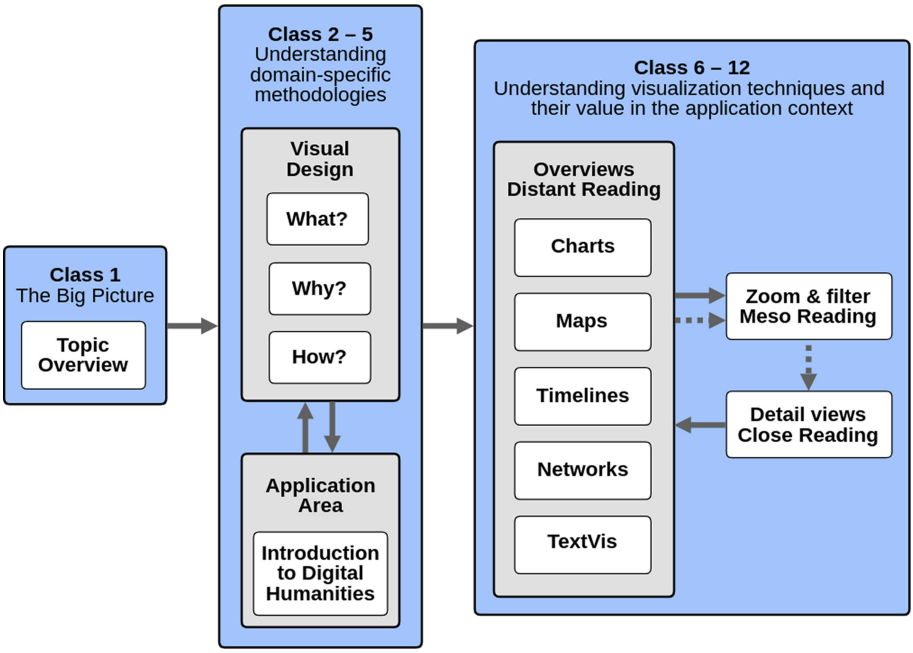

Fig. 4. Suggested structure of the theoretical training.

- Stylometric analysis of a Latin poem collection with graphs [16]
- Quantitative analysis of the Kissinger papers with charts, tag clouds and graphs [35]

3.2 Block 2: Understanding Domain-specific Methodologies

After the students gained an overview of the course topic, they should acquire basic knowledge in the two fields, and they should learn on established research methodologies and process models that lead to knowledge construction.

Application Area. Starting with a concept formation through diverse definitions of digital humanities, we discussed major historical developments that can be seen as the cornerstones of the digital humanities as a research area. One of the early related works, published in 1928, is Wladimir Jakowlewitsch Propp's "Morphology of the Folktale" [47]. Propp analyzed 100 Russian folktales and generated 31 functions, and a subset of which in the given order always defines the plot of a folktale. Another pioneering project is Roberto Busa's "Index Thomisticus". It began in 1948 aiming to create an index of Thomas Aquinas' works that needed to be lemmatized [11]. Recent influential works like Franco Moretti's "Graphs, Maps, Trees" [44], the Culturomics project [43] and Matthew Jocker's "Macroanalysis" [34] lead to defining the term of *distant reading*, which, like no other concept, stands for the importance of visualization for the digital humanities. The collaborative creation of a definition of *digital humanities* and try-out sessions with the Google Books Ngram Viewer[1] are suitable *active learning* tasks.

[1] https://books.google.com/ngrams.

Visual Design. The major goal of the classes related to visual design is that students, independent on their academic backgrounds, are capable to purposefully discuss project ideas using the same terminology. I am engaging in interdisciplinary digital humanities projects for more than ten years. Munzner's nested model for visualization design and validation [45] best reflects our efforts to collaboratively design visualizations that serve their intended purposes [23,27,38]. However, my experiences are even better described by a participatory visual design approach [33], which includes interactive feedback loops between visualization scholars and experts from the targeted domain on all visual design levels. Munzner's book "Visualization Analysis and Design" [46] structures a theoretical information visualization overview closely related to the nested model, thus, it serves as an appropriate basis for this part of the course. The major components are classes on data abstraction to convey *what* data can be visualized, on task abstraction to discuss *why* visualizations are needed, i.e., which user tasks can be supported, and on visual encoding and interaction to overview *how* data can be visualized ensuring that intended tasks are supported. Other visualization text books can be used as a basis for teaching on visual design, e.g., "Information Visualization" books by Ware [61] or Spence [56] both choosing a rather theoretical approach, or Cao and Cui's "Introduction to Text Visualization" [12] for a particular focus on textual data. However, the *domain situation* is an integral component of the nested model, thus, serves best to convey the effectiveness of visualization and the necessity for careful visual design having a particular target domain as the digital humanities at hand.

It is not crucial what research field is discussed first. Introducing the digital humanities as the targeted application domain first makes it easier to refer to the targeted application domain when giving an overview of the visual design. However, my introduction to digital humanities ends with the idea of distant reading, which generates a suitable bridge to specialized distant reading visualizations discussed in the third block of the theoretical training.

Active learning activities should aim to foster the understanding of data and task abstraction. Focusing on a single item or an entire cultural heritage collection, student groups can create appropriate data models, and discuss potential user tasks that could be supported by interactive visualization means. In addition, students should be confronted with existing visualizations aiming to evaluate the appropriateness of their visual mapping. A further suitable activity is to let the students participate in a study to assess the error rates across visual channels [14].

3.3 Block 3: Understanding Visualization Techniques and Their Value in the Application Context

The major part of the theoretical training is dedicated to discussing the most frequently used visualization techniques in the context of digital humanities applications. Each discussed technique basically refers to all visual design levels. Focusing primarily on how data of a specific type can be visualized, attention is given to the tasks that can be supported with a particular visual interface by discussing related published digital humanities applications. Block 3 should be organized on the basis of the Information Seeking Mantra [54], which is applied in many digital humanities applications [28]. However, the digital humanities established its own terminology. The terms distant

reading and marcoanalysis stand for *overview*, meso reading and zooming comply to *zoom and filter*, and close reading and microanalysis relate to *details on-demand*.

Two strategies can be followed by teaching related contents. Either one strictly structures classes according to the Information Seeking Mantra by starting with overview visualizations. However, when students with a humanities background participate, it is more appropriate to begin with a review of visualization techniques that support digital close reading, as they are familiar with traditional close reading techniques. Moreover, it is easier to motivate overview visualizations necessary when operating with a whole collection instead of a single item. For occurring distant reading patterns, means of interaction are required to reach again individual items that are of potential interest for further investigation.

Detail Views. The course on detail views should be motivated by traditional work procedures of humanities scholars who engage with an object of interest. Especially, if students with different study topics participate, this approach contributes to the mutual exchange of domain knowledge and research workflows fruitful for interdisciplinary projects [45]. As textual data is typically focused in my course, I introduced the traditional close reading technique [10,42] and further discussed the benefits of digital environments for close reading tasks [13,36]. Different means of visualizing information and annotating fragments in a close reading view can be discussed in this context. Performing a traditional close reading individually and a collaborative digital close reading are suitable *active learning* tasks. Next to focusing texts, other data types like artworks [7], media contents [3] or individuals [39] can be the basis for discussion.

Zoom and Filter. Exemplarily, I reserved one course to refer to the task to comparatively analyze different text editions on different scales. In this context, I introduced the concept of meso reading explicitly [23]. However, I chose to discuss the majority of means to interactively navigate large data sources such as zooming, panning or focus+context referring to digital humanities projects implementing related means in the courses on specialized overview techniques.

Overviews. The term of distant reading is one of the main pillars of digital humanities research. Related visualization techniques therefore cover the largest share of the course. In the context of related digital humanities projects, the following techniques should be discussed:

- **Charts:** Motivating with historical examples [49], the power of statistical charts like line or bar charts to communicate a distribution of or trend in numerical data should be taught first. Further, grid-based heat map charts, which are frequently used in the literature [28], should be discussed.
- **Maps:** Many applications offer to analyze geographical data using interactive maps. My course informs on the definition of geovisualization in contrast to cartography [40], on different map projections, and on the various geovisualization techniques applied to point- and region based geographical data like choropleth, heat and glyph maps [55].

- **Timelines:** Temporal information is inherent in most data sets digital humanities scholars deal with. Next to referencing the long history of time-based visual representations of information [50], my course includes a discussion of modeling temporal information [1], a review of visualization techniques referring to the space-time cube [2], and a section on uncertainty, which is often comprised in digital humanities data due to the long historical time frame it refers to [8].
- **Graphs:** Relational data is frequently used to generate insights to digital humanities data in a visual form. After an overview of different graph structures to represent relational data, we discussed the basics of drawing different types of graphs [57], e.g., force-based approaches, radial and tree layouts. As standardized tools like Gephi [5] are often applied to generate graphs, the benefits of careful graph design were discussed on the basis of several related works involving visualization scholars [6,19,30].
- **TextVis:** Texts are the most frequently used raw data type in digital humanities applications [28]. Next to basic text processing techniques like part-of-speech tagging and named entity recognition, the course should discuss typical text visualization strategies. First and foremost, this includes traditional [59,60] and adapted tag cloud representations [15,31]. Further, hybrid techniques that adopt the idea of communicating the frequency of words with font size like tag maps [48] or tagged stream graphs [53] are related.

Existing ready-to-use domain-specific visualization frameworks can be used for research-based *active learning* activities. Those may include Google Data Studio[2] for charts, Leaflet[3] or GeoTemCo[4] for geographical data visualization, Timeline.js[5] or Sutori[6] for timeline generation, Gephi[7] or Palladio[8] for graph visualization, and Voyant[9] or TAPoR[10] for visual text analysis. Many of such tools offer multiple views for different types of data. Thus, engaging with them further supports the students' understanding of data modeling and multifaceted visual data exploration. Lastly, it bears also the opportunity to discover and discuss the limitations of existing tools, thereby motivating the development of new, sophisticated solutions.

4 Practical Training

While the theoretical training targets to make students understand the techniques, workflows and research methodologies that lead to knowledge construction, in the practical training students conduct related project tasks on their own. On the one hand, they get

[2] https://datastudio.withgoogle.com/.
[3] https://leafletjs.com/.
[4] http://www.informatik.uni-leipzig.de/geotemco/.
[5] http://timeline.knightlab.com/.
[6] https://www.sutori.com/.
[7] https://gephi.org/.
[8] https://hdlab.stanford.edu/palladio/.
[9] https://voyant-tools.org/.
[10] http://tapor.ca/.

to know how to apply the learned theoretical contents to practical problems, and on the other hand, students need to engage in interdisciplinary collaborations, in which they face issues typical for a digital humanities project.

4.1 Interdisciplinary Projects

Carried out in a *research-based* format, students need to generate solutions for real world inquiries concerning cultural heritage data. Therefore, they need to form interdisciplinary project groups. Dependent on the study backgrounds of the students of a project group, the teacher (or teaching assistant), who needs to be experienced in digital humanities collaborations, needs to be able to take different roles.

The Mediator. The project involves both students having a computer science or a humanities background. In this constellation, carrying out the project based on an research idea generated within the group is entirely in the hands of the students. However, it is especially considerable to participate in project meetings in the role of a mediator being experienced in interdisciplinary collaborations. This ensures that groups stay together in case of crucial issues, for which students lack appropriate problem solution strategies. In such constellations, students undergo the entire life cycle of a digital humanities project (on a small scale) and face typical pitfalls of interdisciplinary projects, and tackling and reflecting on those is the prior targeted skill to be trained in the course.

The Real Humanities Scholar. The project only involves computer sciences students from the course, but they collaborate with an expert humanities scholar to generate a solution to a real world research problem that is of interest to the expert. Depending on the experience of the humanities scholar in digital humanities projects, the role of the course teacher ranges from being the mediator in case of limited to the observer in case of extensive experience in interdisciplinary collaborations. This project setup produced high quality visual design solutions as students could focus on the development while interdisciplinary exchange is better structured as compared to the mediator role projects.

The Fake Humanities Scholar. The project only involves computer sciences students from the course, and the number of project ideas brought by expert humanities scholars is too low. In such cases, I proposed project ideas and took the role of the domain expert in project discussions. However, as I am not trained in humanities subjects, such projects only generate *fake* interdisciplinary discussions, making them the least preferable setup.

The Helper. The project only involves humanities students from the course, and they require technical support to generate appropriate solutions. The role of the teacher includes discussions on data modeling and assistance for generating visual representations of the data. In such constellations, students generated interesting project ideas and produced high-quality data sets being valuable beyond the duration of the course. For visualization purposes, due to limited programming skills, existing tools and libraries were typically applied instead of new solutions being developed.

The project results were presented by the groups in a block seminar at the end of the course. The outcomes of a project in terms of visualization quality and usability played a secondary role in evaluating the results. Reflections on occurring issues concerning the implementation of standardized design models, reported miscommunications among project members, or problems concerning data acquisition and processing were of particular value not only for me as a teacher, but also for the other groups who experienced similar issues. Previous publications on the course report on individual project outcomes in detail [25,32]. The next section is dedicated to those projects that contributed to related publications.

4.2 Publishing Project Results

The research-tutored fragment of the practical training as depicted in Fig. 3 refers to the publishing of project results. This activity was not part of the actual course curriculum, it was rather offered as a training opportunity for scientific writing. The results of three projects were explicitly published.

On the Impact of the Merseburg Incantations [21]. The idea to publish the results of this project existed before, and an already accepted poster presentation at the upcoming annual digital humanities conference 2017 served as a motivation for two students from the humanities department to participate. In the *helper* role, I discussed data modeling and data gathering strategies targeted to collect and structure metadata of published works that, in any form, refer to the Merseburg Incantations. Next to the data set, the students created a geospatial-temporal browser for the collected references based on Google My Maps. A screenshot, which was included in the poster presentation, can be seen in Fig. 5. Unfortunately, both students were not able to reflect on their experiences gained throughout the project at the conference.

Visual Analysis of Engineer's Biographies and Engineering Branches [41]. The project topic was an extension of a Bachelor thesis. The collaborating historian was interested in quantitatively analyzing taught engineering subjects, and in how the engineering branches *materials, manufacturing* and *construction* evolved geospatially and temporally in Germany. While the Bachelor thesis brought forth interactive visualizations to analyze the related prosopographical database of German engineers[11], the course project focused on the development of an interactive tag cloud that supports the assembling of engineering branches. On mouseover, correlations among subjects taught by the same professors are shown, and subjects can be iteratively added to a branch balancing between appropriate and inappropriate co-taught subjects. A screenshot of the tag cloud is shown in Fig. 6. The computer science student carrying out the *real humanities scholar* project presented his results at the LEipzig symposium on Visualization In Applications 2018.

musiXplora: Visual Analysis of a Musicological Encyclopedia [39]. The musiXplora[12] is an online resource for musicological data of different type, e.g., persons related to music history or musical objects. In the course project, a computer science student

[11] https://www.hi.uni-stuttgart.de/gnt/pdm/.
[12] https://home.uni-leipzig.de/mim/musici/.

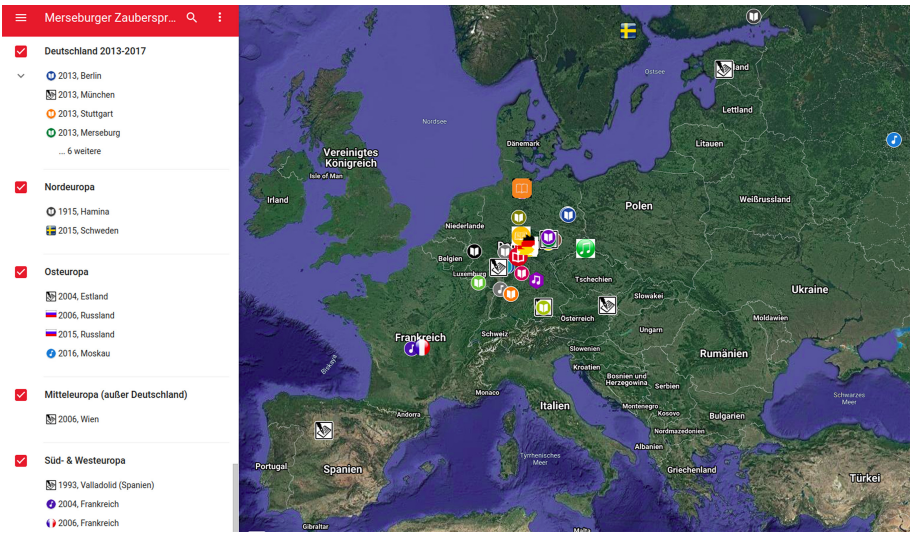

Fig. 5. The published works are grouped according to time and space. Different icons denote different types of works referring to the Merseburg Incantations. Clicking an icon delivers details on the selected work.

developed in close collaboration with Josef Focht, musicology professor and supervisor of the musiXplora project, a design to analyze the career of musical instrument types, e.g., clavichord instruments registered in the musiXplora with different events (e.g., production, purchase, exhibition) attached to them. The visual depiction shown in Fig. 7 reveals the history of the instrument type, i.e., where instruments have been played and when such instruments have been produced. The description of the design was presented by the student at the International Conference on Information Visualization Theory and Applications 2020.

5 Considerations

In the previous version of this paper [32], I discussed different experiences gained in three years teaching the course, including students' reflections on the conducted projects and grading. What follows is a list of considerations valuable for constructing a related course.

Ensure a Diversity of Study Backgrounds! The composition of students mainly influences contents and learning activities of the course. I experienced all possible constellations resulting in the various project roles for me as a teacher. I recommend to offer the course to all interested students from computer science, humanities and digital humanities. This way, methodological exchange among the students is likely and interdisciplinary projects are formed.

Fig. 6. Interactive tag cloud supports assembling engineering branches. The selected subject Wärmelehre (thermodynamics) only co-occurs with subjects already assigned to *construction*, denoted by bold, underlined tokens shown on mouseover.

The More Course Time, the Merrier! The three iterations of my course corresponded to 5 ECTS credits each. This limits not only the contents that can be taught, but also the targeted learning outcomes. Sticking to the theme of the course, the focus in my course was conveying how to design a valuable visualization that supports investigating a real world digital humanities task. This setup did not leave room to exhaustively discuss data transformation procedures necessary to turn raw data sources in a format processible by the visualization. In such a case, one should offer pre-processed data sets and project topics should not require sophisticated data transformation efforts. To convey all aspects of the visual design process for digital humanities, a course subject to 10 ECTS credits is better suitable.

Promote Web-based Developments! JavaScript typically does not belong to the studied programming languages of computer science students, and it is often considered a prerequisite for their studies. However, visualizations used in digital humanities are mostly web-based, and ready-to-use frameworks like D3.js [9] are widely applied. Therefore, I recommend to use widely learned languages such as Python or Java primarily for back end computations, and JavaScript for generating visualizations in a web-based front end. Basics of JavaScript programming could be conveyed in practical sessions parallel to the theoretical training.

Promote Development-driven Visualizations! Digital humanities literature tells us that many visualizations are generated using a standard visualization tool [5]. The generation of arguable visual output in a short amount of time is the main advantage of such ready-to-use frameworks. However, scholars need to learn to operate potentially complex tools and also how to interpret upcoming results. Even more crucial, research interests might deviate from what the tool can provide, which favors new visualization developments. The course should discuss ready-to-use tools, but also point out their limitations.

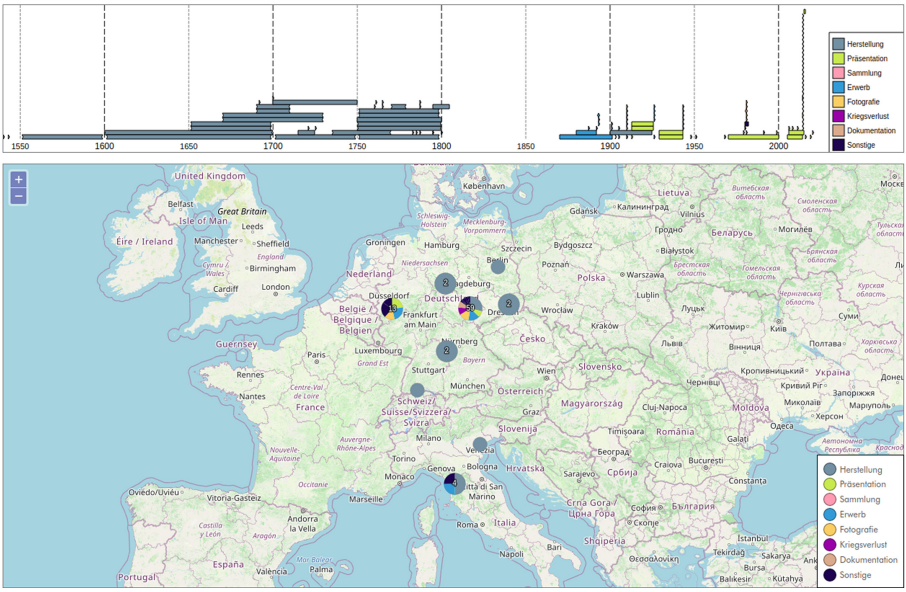

Fig. 7. Geospatial-temporal analysis of the clavichord metadata.

Report on Successes, But, More Importantly, on Failures! As a researcher involved in digital humanities projects, ones own experiences should be discussed thoroughly. This does not only include success stories. Reporting on failures is of particular importance as they document the challenges of interdisciplinary research and diverse merits of generated results. Reporting my failures was important for me being a computer scientist to highlight the necessity to learn more on the research processes and targeted data sources—it is best, if one has a real interest in cultural heritage.

Treat Students Like Researchers! After the theoretical training, students are well equipped with knowledge about visual design and the targeted interdisciplinary research area. Then they are able to implement research ideas in the practical training. Especially participating humanities students generated interesting ideas, and the longer the theoretical training took, the more they used the learned technical terminology to describe their visions. Project discussions, in which the teacher participates, should be perceived and carried out as a meeting among researchers instead of a teacher/student consultation.

6 Conclusion

The insights from my participation in the conducted projects and the students' reflections during project presentations document that a structured research-teaching focusing on understanding and experiencing the research workflows that lead to knowledge construction is worthy for the participating students. While computer science students report on the value to get a detailed overview of a specific application domain for visualization, students of humanities disciplines appreciate to get insights to computational thinking and data modeling making them capable of generating digital humanities

research tasks that could benefit from applying visualization solutions. Independent on their individual study subjects, students seemed to be excited to develop a research idea and to collaboratively design a visualization that supports investigating that idea. Informal evaluations after the course finally certify that the intended goal that students with different backgrounds should learn to "speak the same language" can be reached with the outlined course structure. I hope that the given overview will inspire teachers who aim to teach a similar course on the intersection of visualization and digital humanities.

References

1. Aigner, W., Miksch, S., Schumann, H., Tominski, C.: Visualization of Time-Oriented Data, 1st edn. Springer, Heidelberg (2011). https://doi.org/10.1007/978-0-85729-079-3
2. Bach, B., Dragicevic, P., Archambault, D., Hurter, C., Carpendale, S.: A descriptive framework for temporal data visualizations based on generalized space-time cubes. Comput. Graph. Forum **36**(6), 36–61 (2015)
3. Baker, F.W.: Close Reading and What It Means for Media Literacy (2014). https://www.middleweb.com/15929/close-reading-and-media-literacy/
4. Barker, E., Pelling, C., Bouzarovski, S., Isaksen, L.: Mapping the world of an ancient Greek Historian: the HESTIA project. In: Proceedings of the Digital Humanities 2010 (2010)
5. Bastian, M., Heymann, S., Jacomy, M., et al.: Gephi: an open source software for exploring and manipulating networks. In: ICWSM, vol. 8, pp. 361–362 (2009)
6. Bezerianos, A., Dragicevic, P., Fekete, J., Bae, J., Watson, B.: GeneaQuilts: a system for exploring large genealogies. IEEE Trans. Vis. Comput. Graph. **16**(6), 1073–1081 (2010)
7. Binkis, S.: Reading Visual Images: The First World War (2020). https://dcc.newberry.org/teachers/x5lljd50
8. Börner, K., Eide, O., Mchedlidze, T., Rehbein, M., Scheuermann, G.: Network visualization in the humanities (Dagstuhl Seminar 18482). Dagstuhl Rep. **8**(11), 139–153 (2019). https://doi.org/10.4230/DagRep.8.11.139. http://drops.dagstuhl.de/opus/volltexte/2019/10359
9. Bostock, M., Ogievetsky, V., Heer, J.: D data-driven documents. IEEE Trans. Vis. Comput. Graph. **17**(12), 2301–2309 (2011)
10. Burke, B.: A Close Look at Close Reading: Scaffolding Students with Complex Texts (2012)
11. Busa, R.: The Annals of Humanities Computing: The Index Thomisticus. Raccolta delle opere di Roberto Busa, North Holland Publishing Company (1980). https://books.google.dk/books?id=Ph29oAEACAAJ
12. Cao, N., Cui, W.: Introduction to Text Visualization. Atlantis Briefs in Artificial Intelligence. Atlantis Press, Paris (2016). https://cds.cern.ch/record/2230485
13. Cheema, M.F., Jänicke, S., Scheuermann, G.: AnnotateVis: combining traditional close reading with visual text analysis. In: Workshop on Visualization for the Digital Humanities, IEEE VIS 2016, Baltimore, Maryland, USA (2016)
14. Cleveland, W.S., McGill, R.: The many faces of a scatterplot. J. Am. Stat. Assoc. **79**(388), 807–822 (1984)
15. Collins, C., Viegas, F.B., Wattenberg, M.: Parallel tag clouds to explore and analyze faceted text corpora. In: 2009 IEEE Symposium on Visual Analytics Science and Technology, pp. 91–98, October 2009. https://doi.org/10.1109/VAST.2009.5333443
16. Eder, M.: Stylometry, network analysis, and Latin literature. In: Proceedings of the Digital Humanities 2014 (2014)
17. Healey, M.: Linking research and teaching exploring disciplinary spaces and the role of inquiry-based learning. Reshaping the University: New Relationships between Research, Scholarship and Teaching, pp. 67–78 (2005)

18. Heyer, G., Isemann, D.: Digital und computational humanities. In: Book of Abstracts zum DHd Workshop Informatik und die Digital Humanities (2014)
19. Hinrichs, U., Forlini, S., Moynihan, B.: Speculative practices: utilizing infovis to explore untapped literary collections. IEEE Trans. Vis. Comput. Graph. **22**(1), 429–438 (2016)
20. Hinrichs, U., Forlini, S., Moynihan, B.: In defense of sandcastles: research thinking through visualization in digital humanities. Digit. Scholarsh. Hum. **34**, i80–i99 (2018)
21. Jänicke, S.: On the impact of the Merseburg incantations. In: Conference Abstracts of the Digital Humanities 2017 (2017)
22. Jänicke, S., Franzini, G., Cheema, M.F., Scheuermann, G.: Visual text analysis in digital humanities. Comput. Graph. Forum **36**(6), 226–250 (2017)
23. Jänicke, S., Wrisley, D.J.: Interactive visual alignment of medieval text versions. In: 2017 IEEE Conference on Visual Analytics Science and Technology (VAST), pp. 127–138, October 2017. https://doi.org/10.1109/VAST.2017.8585505
24. Jänicke, S.: Valuable research for visualization and digital humanities: a balancing act. In: Workshop on Visualization for the Digital Humanities, IEEE VIS 2016, Baltimore, Maryland, USA (2016)
25. Jänicke, S.: A visualization course for journalism students. In: IEEE VIS 2019 Workshop on Visualization for Communication (VisComm) (2019)
26. Jänicke, S., Blumenstein, J., Rücker, M., Zeckzer, D., Scheuermann, G.: TagPies: comparative visualization of textual data. In: Proceedings of the 13th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications, pp. 40–51. INSTICC, SciTePress (2018). https://doi.org/10.5220/0006548000400051
27. Jänicke, S., Focht, J., Scheuermann, G.: Interactive visual profiling of musicians. IEEE Trans. Vis. Comput. Graph. **22**(1), 200–209 (2016)
28. Jänicke, S., Franzini, G., Cheema, M.F., Scheuermann, G.: On close and distant reading in digital humanities: a survey and future challenges. In: Borgo, R., Ganovelli, F., Viola, I. (eds.) Eurographics Conference on Visualization (EuroVis) - STARs. The Eurographics Association (2015)
29. Jänicke, S., Gessner, A.: A distant reading visualization for variant graphs. In: Proceedings of the Digital Humanities 2015 (2015)
30. Jänicke, S., Geßner, A., Franzini, G., Terras, M., Mahony, S., Scheuermann, G.: TRAViz: a visualization for variant graphs. Digit. Scholarsh. Hum. **30**(Suppl. 1), i83–i99 (2015)
31. Jänicke, S., Scheuermann, G.: On the visualization of hierarchical relations and tree structures with TagSpheres. In: Braz, J., et al. (eds.) VISIGRAPP 2016. CCIS, vol. 693, pp. 199–219. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-64870-5_10
32. Jänicke, S.: Teaching on the intersection of visualization and digital humanities. In: Proceedings of the 15th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications - Volume 3: IVAPP, pp. 100–109. INSTICC, SciTePress (2020)
33. Jänicke, S., Kaur, P., Kuzmicki, P., Schmidt, J.: Participatory visualization design as an approach to minimize the gap between research and application. In: Gillmann, C., Krone, M., Reina, G., Wischgoll, T. (eds.) VisGap - The Gap between Visualization Research and Visualization Software. The Eurographics Association (2020). https://doi.org/10.2312/visgap.20201108
34. Jockers, M.L.: Macroanalysis: Digital Methods and Literary History, 1st edn. University of Illinois Press, Champaign (2013)
35. Kaufman, M.: 'Everything on paper will be used against me': Quantifying Kissinger. In: Proceedings of the Digital Humanities 2015 (2015)
36. Kehoe, A., Gee, M.: eMargin: a collaborative textual annotation tool. Ariadne (71) (2013)

37. Kerren, A., Stasko, J.T., Dykes, J.: Teaching information visualization. In: Kerren, A., Stasko, J.T., Fekete, J.-D., North, C. (eds.) Information Visualization. LNCS, vol. 4950, pp. 65–91. Springer, Heidelberg (2008). https://doi.org/10.1007/978-3-540-70956-5_4
38. Khulusi, R., Jänicke, S., Kusnick, J., Focht, J.: An interactive chart of biography. In: 2019 IEEE Pacific Visualization Symposium (PacificVis) (2019)
39. Khulusi, R., Kusnick, J., Focht, J., Jänicke, S.: musiXplora: visual analysis of a musicological encyclopedia. In: Proceedings of the 15th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications - Volume 3: IVAPP, pp. 76–87. INSTICC, SciTePress (2020). https://doi.org/10.5220/0008977100760087
40. MacEachren, A.M., Kraak, M.J.: Exploratory cartographic visualization: advancing the agenda (1997)
41. Meinecke, C., Jänicke, S.: Visual analysis of engineer's biographies and engineering branches. In: LEVIA 2018: Leipzig Symposium on Visualization in Applications (2018)
42. Mesmer, H.A., Rose-McCully, M.: A closer look at close reading: three under-the-radar skills needed to comprehend sentences. Read. Teach. **71**(4), 451–461 (2018). https://doi.org/10.1002/trtr.1639. https://ila.onlinelibrary.wiley.com/doi/abs/10.1002/trtr.1639
43. Michel, J.B., et al.: Quantitative analysis of culture using millions of digitized books. Science **331**(6014), 176–182 (2011)
44. Moretti, F.: Graphs, Maps, Trees: Abstract Models for a Literary History. Verso (2005)
45. Munzner, T.: A nested model for visualization design and validation. IEEE Trans. Vis. Comput. Graph. **15**(6), 921–928 (2009). https://doi.org/10.1109/TVCG.2009.111
46. Munzner, T.: Visualization Analysis and Design. CRC Press, Boca Raton (2014)
47. Propp, V.: Morphology of the Folktale, vol. 9. University of Texas Press (2010)
48. Reckziegel, M., Cheema, M.F., Scheuermann, G., Jänicke, S.: Predominance tag maps. IEEE Trans. Vis. Comput. Graph. **24**(6), 1893–1904 (2018)
49. Rendgen, S.: The Minard System: The Complete Statistical Graphics of Charles-Joseph Minard. Chronicle Books (2018)
50. Rosenberg, D., Grafton, A.: Cartographies of Time: A History of the Timeline. Princeton Architectural Press (2013)
51. Rushmeier, H., Dykes, J., Dill, J., Yoon, P.: Revisiting the need for formal education in visualization. IEEE Comput. Graph. Appl. **27**(6), 12–16 (2007)
52. Schwerdtfeger, A.: Erkenntnisgewinn in der Wissenschaft (2020)
53. Shi, L., Wei, F., Liu, S., Tan, L., Lian, X., Zhou, M.: Understanding text corpora with multiple facets. In: 2010 IEEE Symposium on Visual Analytics Science and Technology (VAST), pp. 99–106, October 2010
54. Shneiderman, B.: The eyes have it: a task by data type taxonomy for information visualizations. In: Proceedings of the Visual Languages, pp. 336–343 (1996)
55. Slocum, T.A., McMaster, R.B., Kessler, F.C., Howard, H.H.: Thematic Cartography and Geovisualization. Prentice Hall Series in Geographic Information Science, 3rd international edn. Prentice Hall (2009)
56. Spence, R.: Information Visualization, vol. 1. Springer, Heidelberg (2001)
57. Tamassia, R.: Handbook of Graph Drawing and Visualization (Discrete Mathematics and Its Applications). Chapman & Hall/CRC (2007)
58. Tominski, C., Schumann, H.: Interactive Visual Data Analysis. AK Peters Visualization Series, CRC Press (2020). https://doi.org/10.1201/9781315152707. https://ivda-book.de
59. Viegas, F., Wattenberg, M., Feinberg, J.: Participatory visualization with wordle. IEEE Trans. Vis. Comput. Graph. **15**(6), 1137–1144 (2009)
60. Viegas, F., Wattenberg, M., van Ham, F., Kriss, J., McKeon, M.: ManyEyes: a site for visualization at internet scale. IEEE Trans. Vis. Comput. Graph. **13**(6), 1121–1128 (2007)
61. Ware, C.: Information Visualization: Perception for Design. Elsevier (2012)
62. Windhager, F., et al.: Visualization of cultural heritage collection data: state of the art and future challenges. IEEE Trans. Vis. Comput. Graph. **25**(6), 2311–2330 (2019)