Chapter 8
Conclusion

> All the pieces are there – huge amounts of information, a great need to clearly and accurately portray them, and the physical means for doing so. What has been lacking is a broad understanding of how best to do it.
>
> Wainer (1997, p. 112)

This book has dealt with concepts and methods for visualizing time and time-oriented data. In this chapter, we will briefly summarize the key aspects that have been elaborated so far. We will further consider the current state of visualization of time-oriented data and identify application issues and research challenges. We conclude with the description of a blueprint of concepts that may serve as a basis for tackling the identified problems in the future.

8.1 Summary

Most of the existing techniques for computational analysis and visualization have been developed to deal with numbers. In practice, these numbers are often anchored in space and time. Depictions of a spatial frame of reference lie at the core of cartography and geo-visualization, and these disciplines have yielded many powerful visualization techniques. However, until now, no independent research field which focuses on the visualization of time-oriented data has been established. Nevertheless, visual representations of time and time-oriented data have a long and venerable history. This fact has been illustrated by means of several classic examples from the pre-computer era in Chapter 2.

Designing appropriate visual representations for time-oriented data requires responding to the questions:

- *What* is presented?
- *Why* is it presented?
- *How* is it presented?

W. Aigner et al., *Visualization of Time-Oriented Data*,
Human-Computer Interaction Series, DOI 10.1007/978-0-85729-079-3_8,

So, in order to adequately address the *what* aspect, one needs to consider the specific characteristics of the time domain as well as the characteristics of the data that are related to time. In Chapter 3, we used four principal criteria to categorize the characteristics of time: the scale (ordinal vs. discrete vs. continuous), the scope (linear vs. cyclic), the arrangement (point-based vs. interval-based), and the viewpoint (linear vs. branching vs. multiple perspectives). To characterize the data, four major criteria have been used: the scale of variables (qualitative vs. quantitative), the frame of reference (abstract vs. spatial), the kind of data (events vs. states), and the dimensionality (univariate vs. multivariate). We further defined time primitives that act as a kind of glue between time and data. Time primitives are often organized in hierarchically structured calendar systems to accommodate different levels of temporal granularity (e.g., seconds, minutes, hours).

With the *why* question user tasks come into play. We addressed this in Chapter 4. At the task level we distinguished lookup and comparison tasks, direct and indirect search, as well as elementary (performed on individual values) and synoptic tasks (performed on sets of values). As a matter of fact, the particular characteristics of time and data as well as the specific user tasks largely influence the design of visualization solutions.

General principles of *how* time and time-oriented data can be visualized were presented in Chapter 4 as well. Two basic distinctions were made: the dimension of time can be represented using the display space (i.e., static representation) or physical time (i.e., dynamic representation), where the presentation space can be either two-dimensional or three-dimensional. We discussed these categories in detail and gave various examples of visualization designs addressing specific aspects of the data level, the task level, and the presentation level.

Visual exploration and analysis of time-oriented data also requires interaction methods that allow users to manipulate the visual representation in a variety of ways, including navigation of time and data, adjustment of graphical encoding and spatial arrangement, selection of data of interest, filtering out irrelevant data, and many more. Chapter 5 provided a compact overview of interaction.

Moreover, analytical methods have to be provided for supporting the generation of expressive visual representations. Among other purposes, analytical methods are useful for computing data abstractions that may serve to cope with large volumes of data or to allow visual analysis at different levels of granularity. Chapter 6 was dedicated to the aspect of analytical support.

As diverse and manifold as time and data characteristics and visualization design choices, interaction concepts, and analytical methods are, the body of available visualization techniques for time and time-oriented data is equally diverse. In Chapter 7, we summarized many classic and state-of-the-art techniques, some of a general nature, others very specific, and some well-established, and others more progressive or alternative. We categorized all of these techniques according to six major criteria, described briefly the main idea behind the techniques, and provided illustration examples. The result is a compact overview of the variety of existing visual representations for time and time-oriented data.

In conclusion, the survey gives evidence that time is indeed an important dimension that deserves special treatment in visual, interactive, and analytical methods. However, tools or systems that provide the broad functionality demanded specifically for time-oriented data are not available. Moreover, there are several open issues to be addressed in the future. In the next sections, we will take a look at these issues from an application perspective as well as from a research point of view.

8.2 Application Issues

A main concern from an application perspective is to bridge the gap between the development of powerful visual methods on the one hand, and their integration into the real-life workflows in different application scenarios on the other hand. This requires research which addresses the problems in terms of both software issues and user issues.

Software issues – systems & formats

There are a variety of commercial and open source visualization systems, as for instance, Tableau[1], Spotfire[2], MagnaView[3], or vtk[4]. Many of the available systems provide excellent support for visual exploration and analysis of multivariate data. However, the specifics of time are not always easy to handle, because time is treated as just another quantitative variable or because the software lacks support for the wide range of characteristics which are relevant when dealing with time (e.g., support for cyclic time or for different time primitives). As a result, in order to visually explore temporal dependencies, the user has to manually find an appropriate visual mapping that emphasizes the time axis. Moreover, this means it is difficult or actually not feasible for users to apply a particular system for all the different characteristics of time discussed before.

On the other hand, visualization research has yielded powerful research prototypes that provide dedicated support for the time aspect. A prominent example in this regard is the TimeSearcher[5] project for visual exploration of time-series data (↪ p. 188). However, the integration of such prototypes into the infrastructure of day-to-day business is usually problematic and requires additional effort. Furthermore, research prototypes are usually not designed to cover all aspects of time, but instead to address only particular cases – mostly the visualization of linear and ordered time domains.

1 http://www.tableausoftware.com/; Retrieved Feb., 2011.

2 http://spotfire.tibco.com/; Retrieved Feb., 2011.

3 http://www.magnaview.nl/; Retrieved Feb., 2011.

4 http://www.vtk.org/; Retrieved Feb., 2011.

5 http://www.cs.umd.edu/hcil/timesearcher/; Retrieved Feb., 2011.

Another significant problem to be solved is caused by the diversity of existing data formats and interfaces. Processes that generate or collect data and tools that manipulate or analyze the data often use specific databases and data formats that meet the requirements of the particular application scenario. Software tools for visualizing the data and interacting with them often use different formats. This circumstance requires individual and possibly complex data transformations, which represent a substantial obstacle. To this end, researchers who develop more comprehensible and simplified data interfaces will be rewarded with greatly expanded user communities.

User issues – knowledge & support

Besides improving the technical basis, it is important to take the needs of the users into account. In this regard, an important point is to improve the awareness about new visualization and interaction methods. Nowadays, users mostly apply traditional visualization techniques such as line plots or bar graphs. These techniques are well-established and have proven to be useful. However, new innovative visualization methods allow the representation of a larger number of variables and data values, provide comprehensive interaction functionality, and take the specific aspects of time into account. These new possibilities can lead to new findings.

Moreover, in many application domains, visual methods are primarily used to present results. This is an adequate strategy for all those analytical problems whose solution can be computed automatically and that do not require input from the user to generate analysis results. However, many analytical problems do not have a single closed solution, but rather require an interactive visual exploration process that integrates the user tightly throughout the analysis workflow.

Another point to mention is that typically users in specific application domains are the ones to create visual representations. This implies that they have to know which visual representation should be used for which task. However, although these users have a strong domain-specific background, it can not be assumed that they are also experts in visualization design. If users were better supported by the visualization systems in choosing expressive, effective, and appropriate visualization techniques, the quality of information display and analysis results could greatly improve. To this end, facilities for specifying data characteristics and analytic tasks are mandatory.

Moreover, cumbersome data transformations and extensive switching between application and visualization systems are substantial obstacles for widespread use. In fact, interactive visualization methods for time-oriented data have to be integrated into application portals and systems in order to allow domain experts to use these techniques effortlessly and seamlessly.

To summarize, bridging the gap between research on interactive visualization methods and their application requires both imparting an awareness of the variety of possibilities and providing means to effectively use them within a given application infrastructure.

8.3 Research Challenges

Designing appropriate visual representations and tools for time-oriented data as well as making them applicable in real world problem solving scenarios requires further scientific investigation. In the following, we will discuss several aspects in this regard.

Problem specification & user guidance

In the field of software engineering it is generally acknowledged that the first step in developing tools and user interfaces should be a sound analysis of the given problem domain (see Hackos and Redish, 1998; Courage and Baxter, 2005). The same applies for designing visual representations. However, most of today's visualization systems do not provide any means of describing the visualization problem. In the case of visually analyzing time-oriented data, this means (1) the characteristics of time and associated time-oriented data as well as (2) the intentions and tasks of users have to be specified. If appropriate descriptions are provided to store this knowledge, visualizations can be realized that suit the data and the tasks. Although first approaches for automatic visualization design have been developed (see Mackinlay, 1986; Senay and Ignatius, 1994; Wills and Wilkinson, 2010), further research is needed to allow an easy-to-use specification of data and tasks and the provision of adequate descriptors, in this way enabling the automatic suggestion and computation of appropriate visual representations. The aim is to guide the users, rather than burden them with technical details. Thus, a significant shift could be realized from a technique-centered view to a user-centered view, i.e., a view that puts the user into the focus (see Kerren et al., 2007).

New visualization methods

Choosing adequate visualizations is one point, but providing a set of suitable techniques that cover all the different aspects of time is another concern. Although a large diversity of powerful visualization techniques for time-oriented data have been developed, most of them support only certain parts of the introduced time and data categorization. In the particular case of visualizing multivariate data, usually linear, point-based, and ordered time domains are assumed. Further investigations are required, including the development of techniques for interval-based time, branching time, and multiple perspectives, for simultaneously displaying raw data and data abstractions, and for showing the time-oriented data in their spatial frame of reference.

Another important challenge to be considered originates from the hierarchical nature of time, which leads to the situation that data may be given at multiple levels of granularity. New interactive visualization techniques are required to allow analysts to combine different levels of data and time and to switch between the levels.

This specific challenge for time-oriented data is related to the more general open research topic of handling variables given at multiple scales (see Keim et al., 2010).

Data quality & data provenance

We have primarily considered visual representations that show the data themselves, rather than information on data quality or data provenance. However, taking uncertainties and provenance into account will significantly improve the expressiveness of visual representations, and also strengthen the user's confidence in the findings made. In the case of time-oriented data, uncertainties of the data and uncertainties of the temporal frame of reference have to be communicated. Currently, visualization methods communicating both time and data uncertainties are not available. Therefore, new visualization strategies have to be developed. The concern of representing the quality of data has also been acknowledged as an open research topic in visualization research in general (see Keim et al., 2010).

Scalability

Apart from issues concerning data quality, the fast growing quantities of data also pose a considerable challenge to visualization. Not too long ago researchers and practitioners were exploring and analyzing acquired or simulated datasets of modest size (e.g., from kilobytes to a few megabytes). Nowadays, it is common to focus on problems involving enormous amounts of data in gigabytes, terabytes, petabytes, or more (see Ward et al., 2010). This calls for particular characteristics and functionalities of data management as well as of the visual and algorithmic design (see Keim et al., 2010). With regard to time-oriented data, we need to be able to deal with both very long time-series containing vast amounts of time primitives (e.g., covering large time spans and/or at very fine-grained time scales) and large numbers of time-dependent variables in parallel. If the visualization and interaction techniques as well as the analytical methods are able to cope with such large quantities of time-oriented data, we will be able to extend the frontiers of complex and demanding research areas, such as biomedicine or climate research.

Novel interaction methods

Interaction methods are essential to allow users to explore time-oriented data as well as the space of possible visual encodings. Obviously, the features of time-oriented data influence the interaction techniques and tasks carried out on such data. Navigating in time and switching between different levels of temporal granularity are mandatory when interacting with time-oriented data, but are rather uncommon when interacting with abstract quantitative variables. The visualization of larger datasets benefits from visual overviews and the ability to drill down into areas of interest while preserving orientation within the information space. Interacting directly with visual representations and analytical methods provides more control and tighter feedback for the human analyst.

However, interaction techniques for time-oriented data mostly consider linear, point-based time domains. More research is needed to enable users to interact with visual representation of cyclic time, branching time, or multiple perspectives. This also includes reasoning about what it means to move in cycles, along branches, or in multiple perspectives. Moreover, support is needed to structure the interactive exploration process and to provide guidance in terms of where to go next or which encoding to choose. It is important that these investigations be made in accordance with the users' demands. A recent workshop on "Interacting with Temporal Data" organized at the ACM CHI Conference 2009 gives evidence of the growing interest in this research topic (see Mackay et al., 2009).

Advanced analytical methods

Most analytical methods for time-oriented data available today treat time as a flat, ordered sequence of events. Thus, these methods are lacking information about the time intervals between events or about after how much time a particular pattern will reoccur. But, the natural structures of time, such as years and seasons, as well as social structures of time, such as weeks and business days, can strongly influence what findings can be extracted from time-oriented data. For example, the pattern of monthly sales may vary largely due to differences in the arrangements of workdays, weekends, and holidays. However, only few existing analytical methods, like for example the seasonally adjusted autoregressive integrated moving average (SARIMA), model cyclic temporal behavior adequately. As a consequence, better support for dealing with the hierarchical and cyclical structures of time as well as their semantics (i.e., accounting for the specifics of time) is needed.

Moreover, many analytical methods are like black-boxes which accept some time-oriented data as input and generate some analytic result as output. However, it remains largely unexplored how to parameterize analytical methods appropriately to adapt to the given data and tasks, or how to combine multiple analytical methods to generate better analysis results. To solve these problems, the black-boxes must be made transparent and steerable for the user, where sufficient support by the visualization system should be taken for granted.

Evaluation

In order to be able to automatically suggest techniques to the user or to automatically adapt chosen techniques to data and tasks, we need to know which techniques are *good*. This requires evaluation. Evaluation has to be conducted in terms of the three criteria expressiveness, effectiveness, and appropriateness (see Chapter 1). Expressiveness and effectiveness are related to the data level and the task level, respectively. They require testing whether the characteristics of time and data are sufficiently communicated, and whether the visual representation matches the tasks, expectations, and perceptual capabilities of users. With the appropriateness criterion, resources come into play. Moreover, the application domain has to be taken into account to address the appropriateness criterion.

Because thorough evaluation requires a combined consideration of multiple criteria, Munzner (2009) introduced a nested model for visualization design and validation. She proposed to subdivide the generation of visual representations into four nested levels (i.e., characterization of data and tasks, abstraction into operation and data types, design of encoding and interaction, and development of algorithms) and argues that distinct evaluation methodologies should be used for each level of the model. Plaisant (2004) gives an overview of currently applied evaluation methods and points out challenges specific to the evaluation of information visualization. These challenges also apply to evaluating time-oriented data visualizations. Although general evaluation methods (see Lazar et al., 2010) and methods tailored to visualization, for example, to measure effectiveness (see Zhu, 2007), are readily available, more research is needed in terms of specifically evaluating the encoding of and the interaction with time and time-oriented data. To this end, it is necessary to develop new evaluation strategies that are tailored to the requirements of visualizing time-oriented data.

Multi-user & heterogeneous display environments

Typically, visualization and interaction techniques are designed to be used in single user and single display scenarios. However, modern technologies and infrastructures enable analysts to use multi-touch displays with tangible interaction, to operate large-display environments with virtual interaction, or to work in smart environments, which are heterogeneous conglomerates of dynamically linked devices. On the one hand, technological progress opens up new possibilities. For example, one can imagine multi-user and multi-display problem solving scenarios, where a large-scale display shows an overview of a large time-oriented dataset, while multiple user groups discuss selected details from different parts of the dataset shown on multiple linked tabletop displays. On the other hand, new research questions and challenges have to be addressed. From a technical point of view, there is a need to scale visualization, interaction, and analytical facilities to the given resources (display, interaction, and computation devices) and to the intended audience (single user, small or large user groups). From a user perspective, research is necessary to investigate how multiple users can work together to analyze time-oriented data cooperatively. On a more general level, these issues are related to the challenges of display scalability and human scalability as described by Thomas and Cook (2005).

Non-visual mapping and accessibility

Apart from the numerous options of visually representing time-oriented data we have seen throughout the book, other forms of representing time are possible. Time could for example be mapped to sound or to haptic sensations as with braille interfaces. Smell and flavor might also be candidates for such a mapping. Despite the fact that these mappings are in principle imaginable, their feasibility and usefulness have to be investigated. Particularly, this also addresses accessibility by users with disabilities, especially blind users. Speeth (1961) already showed how seismographic data can be presented in an auditory display. An example of a recent attempt in this direction is a system for data sonification by Zhao et al. (2008) to explore spatial data for users with visual impairments. Apart from being an essential issue for government agencies this is also a fruitful research topic that can lead to novel insights.

Intertwining visual, interactive, and analytical methods

Finally, there is a challenge that touches – or better, that encompasses all previously mentioned concerns: In order to facilitate exploration and analysis of time-oriented data, we should not consider visual, interactive, and analytical methods in isolation, but instead should strive for a tight intertwining of them, effectively utilizing their strengths and compensating their weak spots. In the next section, we give an outlook in this direction.

8.4 Visual Analytics

The development of new methods is only one side of the coin; combining them efficiently and integrating them into the workflow of the user is the other side. The new research field of *visual analytics* aims to combine visual, interactive, and analytical methods within real scenarios to solve real application problems (see Thomas and Cook, 2005; Keim et al., 2010).

A basic precondition for achieving this goal is to provide an open, extensible framework that:

- comes with interchangeable building blocks for visualization, interaction, and analysis of time-oriented data,
- uses commonly agreed-upon interfaces for flexible combination of methods,
- provides descriptors for specifying available techniques as well as data and user tasks, and
- offers an easy-to-use user interface for interactive exploration, including on-demand guidance.

Designing such a comprehensive framework is a formidable research challenge, not to mention the effort required to actually implement the framework's broad functionality. To make a very first step in this direction, we sketch a basic design concept in Figure 8.1.

Data component The data component encapsulates all data-related aspects, including data import from heterogeneous sources, flexible data management, and efficient data search and retrieval. Such functionality is typically realized by database management systems (DBMS), where the underlying paradigm can be manifold (e.g., relational model, temporal database, key-value store, column store, graph database, etc.).

The data component contains time-oriented data, which typically come from various external data sources in different idiosyncratic formats. DBMS functionality can be used to consolidate, unify, and combine the different data sources.

In order to support the seamless interplay of all of the components of the framework, appropriate specifications are needed. This is reflected by the descriptor database. This database includes metadata about the data under investigation such as characteristics of the time domain and characteristics of associated time-oriented data (see Chapter 3).

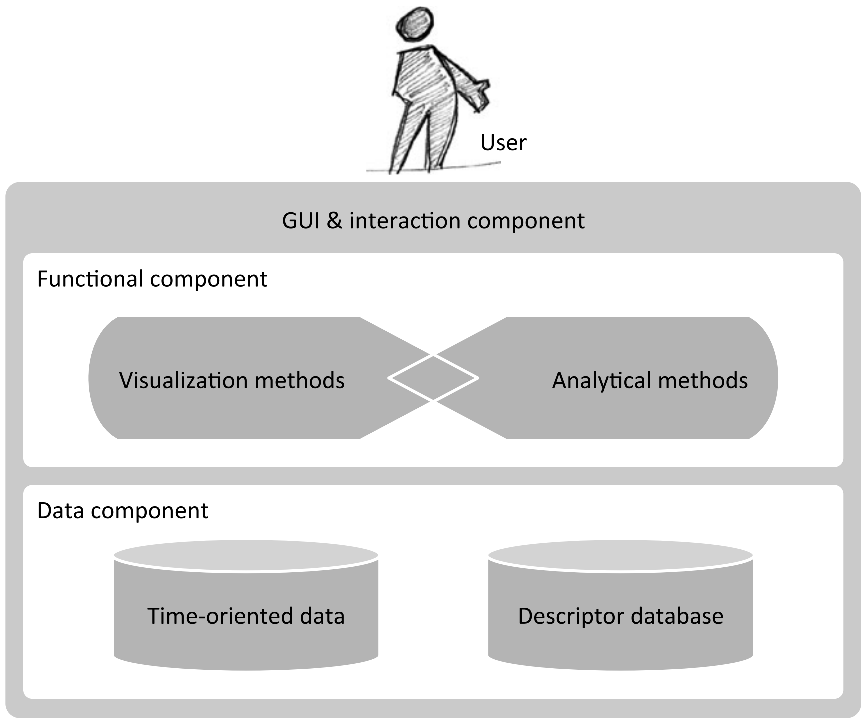

Fig. 8.1: Conceptual framework for visual analytics of time-oriented data.

Besides the mandatory data specification, the descriptor database should provide further useful information such as a description of user tasks and a specification of the capabilities of visual, interactive, and analytical methods. Unfortunately, these latter aspects are rarely supported by current systems. But only if data, tasks, and capabilities are properly specified is it possible to automatically suggest suitable (combinations of) visualization, interaction, and analysis methods for any particular combination of data and tasks. An example for the potential of automatic adaptation according to some specification is the concept of event-based visualization as described in Section 5.4.

An interesting question is how to collect the information to fill the descriptor database. Although some information can be generated automatically, for example, by computing statistics such as minimum, maximum, or data distribution, in most cases it is necessary to resort to manual specification. Therefore, the framework should contain support for both automatic processing to create descriptors as well as interactive specification and refinement to complete the descriptors. Furthermore, there is a need to change descriptors dynamically during the visual analysis process. For example, analysis results may lead to new hypotheses about the data, which in turn manifest themselves in the formulation of new user tasks, in a switch of interests, or in the need to access additional data, which affects the generation of metadata.

Functional component The heart of the framework is given by the functional component that includes visualization methods and analytical methods. The set of visualization methods can be filled with the techniques described in Chapter 7. The set of analytical methods includes computational approaches for analyzing and mining time-oriented data as described in Chapter 6.

GUI & interaction component The third and very important component consists of an all-encompassing graphical user interface and corresponding interaction methods. This component provides the means for embedding visual analytics in day-to-day workflows and for allowing for interactive exploration and analysis of the time-oriented data as discussed in Chapter 5.

In the previous two paragraphs we used only a few lines to discuss the functional component (visualization and analysis) and the GUI & interaction component, because these aspects have been detailed in separate chapters in this book. Notably, there are many different techniques available, but they are considered in isolation. The big challenge of visual analytics is to facilitate smooth interoperability among different approaches to the exploration and analysis of time-oriented data.

If we arrive at commonly agreed-upon interfaces, the building blocks and methods of the framework can be integrated into application portals and thus into the workflows of users. If we succeed in providing appropriate descriptors, automatic selection of suitable techniques and corresponding configurations can be accomplished, effectively guiding the user throughout the visual analysis process. This leads to the new paradigm of making the users and their tasks the focal point, where a central aspect is to assemble visual analytics tools automatically as illustrated in

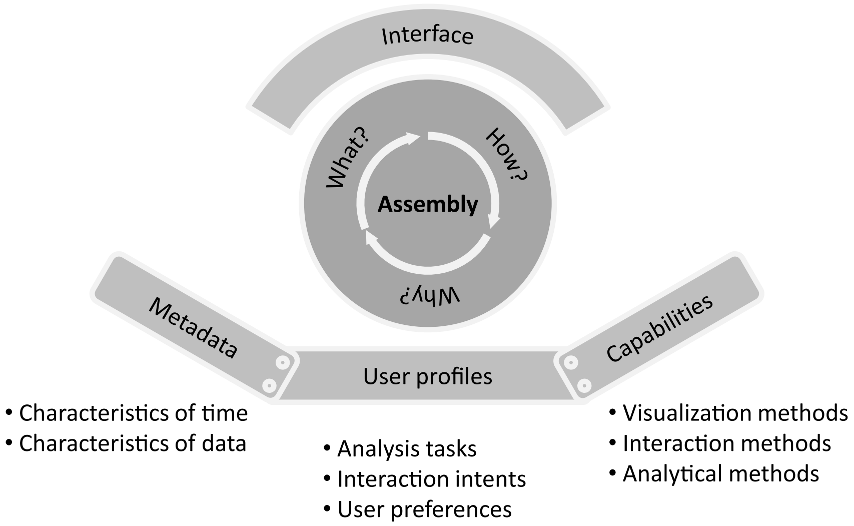

Fig. 8.2: Automatic assembly of visual analytics components.

Figure 8.2. This automatic approach aims to relieve users of the burden of compiling solutions manually. As a result, users can focus on accomplishing tasks on the data by utilizing the power of analytical methods to verify hypotheses and the flexibility of interactive exploration to discover unexpected findings.

Implementing such a user-centric view is one of the most challenging problems identified by the current research agenda of visual analytics (see Keim et al., 2010). First steps have been taken to solve this problem and contributing to an improved analysis of time-oriented data is one of them.

References

Courage, C. and Baxter, K. (2005). *Understanding Your Users*. Morgan Kaufmann, San Francisco, CA, USA.

Hackos, J. T. and Redish, J. C. (1998). *User and Task Analysis for Interface Design*. John Wiley & Sons, Inc., New York, NY, USA.

Keim, D., Kohlhammer, J., Ellis, G., and Mansmann, F., editors (2010). *Mastering the Information Age – Solving Problems with Visual Analytics*. Eurographics Association, Geneve, Switzerland.

Kerren, A., Ebert, A., and Meyer, J., editors (2007). *Human-Centered Visualization Environments*, volume 4417 of *Lecture Notes in Computer Science*. Springer, Berlin, Germany.

Lazar, J., Feng, J. H., and Hochheiser, H. (2010). *Research Methods in Human-Computer Interaction*. John Wiley & Sons, Ltd., London, UK.

Mackay, W. E., Van Kleek, M. G., and Tabard, A. (2009). Interacting with Temporal Data. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI)*, pages 4783–4786, New York, NY, USA. ACM Press. Extended Abstracts.

Mackinlay, J. (1986). Automating the Design of Graphical Presentations of Relational Information. *ACM Transactions on Graphics*, 5(2):110–141.

Munzner, T. (2009). A Nested Process Model for Visualization Design and Validation. *IEEE Transactions on Visualization and Computer Graphics*, 15(6):921–928.

Plaisant, C. (2004). The Callenge of Information Visualization Evaluation. In *Proceedings of the Working Conference on Advanced Visual Interfaces (AVI)*, pages 106–119, New York, NY, USA. ACM Press.

Senay, H. and Ignatius, E. (1994). A Knowledge-Based System for Visualization Design. *Computer Graphics and Applications*, 14(6):36–47.

Speeth, S. D. (1961). Seismometer Sounds. *The Journal of the Acoustical Society of America*, 33(7):909–916.

Thomas, J. J. and Cook, K. A. (2005). *Illuminating the Path: The Research and Development Agenda for Visual Analytics*. IEEE Computer Society, Los Alamitos, CA, USA.

Wainer, H. (1997). *Visual Revelations: Graphical Tales of Fate and Deception from Napoleon Bonaparte to Ross Perot*. Copernicus, New York, NY, USA.

Ward, M., Grinstein, G., and Keim, D. (2010). *Interactive Data Visualization: Foundations, Techniques, and Applications*. A K Peters Ltd, Natick, MA, USA.

Wills, G. and Wilkinson, L. (2010). AutoVis: Automatic Visualization. *Information Visualization*, 9(1):47–69.

Zhao, H., Plaisant, C., Shneiderman, B., and Lazar, J. (2008). Data Sonification for Users with Visual Impairment: A Case Study with Georeferenced Data. *ACM Transactions on Computer-Human Interaction*, 15(1):4:1–4:28.

Zhu, Y. (2007). Measuring Effective Data Visualization. In *Proceedings of the International Symposium on Visual Computing (ISVC)*, pages 652–661, Berlin, Germany. Springer.