Visual Analytics Meets Process Mining: Challenges and Opportunities

Theresia Gschwandtner

Vienna University of Technology,
CVAST – Centre for Visual Analytics Science and Technology
gschwandtner@cvast.ac.at
http://www.cvast.tuwien.ac.at/cvast

Abstract. Event data or traces of activities often exhibit unexpected behavior and complex relations. Thus, before and during the application of automated analysis methods, such as process mining algorithms, the analyst needs to investigate and understand the data at hand in order to decide which analysis methods might be appropriate. Visual analytics integrates the outstanding capabilities of humans in terms of visual information exploration with the enormous processing power of computers to form a powerful knowledge discovery environment. The combination of visual data exploration with process mining algorithms makes complex information structures more comprehensible and facilitates new insights. In this position paper I portray various concepts of interactive visual support for process mining, focusing on the challenges, but also the great opportunities for analyzing process data with visual analytics methods.

Keywords: Visual Process Mining, Visual Analytics, Challenges

1 Introduction

Today we are confronted with an overload of data and information. The amount of information produced on a daily basis is just too much for anyone to process in order to answer specific questions. Examples include data from electronic health records, real time sensors, communication logs, and financial transactions. However, somewhere within this huge amounts of data, there is valuable information to solve specific tasks and answer specific questions. Yet, in order to make sense of the data, we need to find appropriate ways to process it. There are two opposed approaches to tackle this problem: (1) machine learning and data mining, and (2) visualization of the data.

The first approach, i.e. using machine learning and data mining, taps the computational power of the computer for statistical analysis of the data. As a result, a report is generated, giving - at best - the answer to the user's question. However, the second approach, i.e. visualizing the data, utilizes the power of human perception to simultaneously process large amounts of data. When visualizations are carefully designed (e.g., using a suitable combination of preattentive visual attributes to encode important information), the human perception is capable of processing large amounts of data in parallel and of quickly identifying patterns and outliers [34]. Thus, visualization helps to provide an overview of large and complex data. Furthermore, interactive controls allow the user to explore different aspects of the data in more detail.

Visual analytics is defined as 'the science of analytical reasoning facilitated by visual interactive interfaces' [45, p. 4]. It exploits both, the computational power of the computer and the human's perception system to facilitate insights and enable knowledge discovery in large and complex bodies of data. Thus, visual analytics combines both of the aforementioned approaches: machine learning and data mining, and visualization of the data.

1.1 Benefits of Visualization

One example that motivates the use of visualizations nicely, is the Anscombe's Quartet: In 1973 the statistician Francis Anscombe constructed four data sets [5] to demonstrate the necessity to analyze data with visual means as well as the impact of outliers on statistical properties. The summary statistic properties of these four data sets are nearly identical which could lead to the conclusion that these data sets are nearly identical. However, visualizing these data sets with scatter-plots reveals very different behavior of all four sets.

Another example is given by Spence and Garrison [44]: The Hertzsprung Russel Diagram shows the temperature of stars plotted against their magnitude. When looking at this diagram humans can easily distinguish different clusters of types of stars, while automatic means fail due to the noise and artifacts within the data. Card et al. summarize the ways is which information visualization amplifies human cognition like this [10]:

– Increasing working resources such as by using a visual resource to offload work from the cognitive to the perceptional system
– Reducing search such as by representing large amounts of data and grouping information
– Enhancing the recognition of patterns such as by visually organizing data by structural relationships
– Supporting perceptual inference of relationships by suiting the human's perceptual abilities
– Perceptual monitoring of a large number of potential events such as by making changes stand out visually
– Providing a manipulable medium that, unlike static diagrams, allows for interactively exploring the data

In contrast to pure question answering tools, such as IBM Watson [13] (a computing system that uses machine learning technologies to answer specific questions), visual analytics combines these technologies with interactive presentations of data and results, and thus, it enables users to derive additional insights from interacting with the data. Hence, users are able to browse the data and they may find answers to questions they did not even know they were looking for.

2 Visual Approaches for Process Mining

In their process mining manifesto [47] van der Aalst et al. emphasize the potential of visual analytics, i.e. the combination of visualizations, interactions, and mining techniques, to enhance process mining. We conducted a literature research starting with the process mining manifesto and references to relevant systems we already knew. Moreover, we used libraries and search engines such as Google Scholar [16] and IEEE Xplore [24] to search for various related keywords and we went through proceedings of important conferences such as IEEE VIS [23], EuroVis [12] conference, and the BPM conference [11], as well as through the list of references of relevant papers.

Most process mining approaches employ some kind of visual representation due to the wealth and complexity of the data. Yet, they usually do not explicitly consider visualization as a key factor and rather put emphasis on the mining methods [46, 47]. Visualizations are used, for instance, for process discovery: to explore the event log data and to understand, reason about, and fine-tune the derived models (e.g. [26, 20, 51, 32, 22, 48, 21]). To this end, flow charts [36] and directed graphs are used most frequently. Other examples use visualizations to identify interesting behavior and patterns in the event log data (e.g., [43, 7]), or to investigate process conformance (e.g. [49, 1, 8]). In this position paper I give by no means a comprehensive list of visualization approaches in the context of process mining. Yet, I introduce a selection of our own approaches, highlighting some alternative visual representations and interaction techniques in order to spark ideas how to combine process mining with visual analytics.

2.1 Visual Analysis of Guideline Conformance

A related topic to the analysis of process conformance is the analysis of conformance of clinical actions with the recommended care process. Clinical practice guidelines give specific recommendations on how to treat a patient in a given clinical situation. They are aimed at assisting the care process and ensuring its quality [14]. Hence, they can be seen as process models. Bodesinsky et al. [8] presented a visual interactive approach to analyze the compliance of clinical care with these guidelines. The approach is aimed at supporting (1) physicians at the point of care to display clinical guidelines and point them to omitted clinical actions, and (2) medical experts in retrospective analysis of the quality of the treatment and also of the quality of the clinical practice guideline. It visualizes the recommended sequence of actions (see Figure 1 (a)) together with the actual clinical actions that were applied (see Figure 1 (c)). Applied clinical actions are represented by diamonds on a time axis (a pop-up window gives the user details about these actions when hovering them). Moreover, the patient's parameters (see Figure 1 (b)) and the compliance of actions with the clinical guidelines (see Figure 1 (c)) are represented on a time axis. To aid the analysis of conformance, they visually distinguish conform and non-conform actions, as well as missing actions. Conform actions are indicated by diamonds and non-conform actions are indicated by diamonds marked with an X. Any delay of application is indicated by two parallel colored bars, marking the time span between the intended time of application and the actual time of application (see Figure 1 (c)).

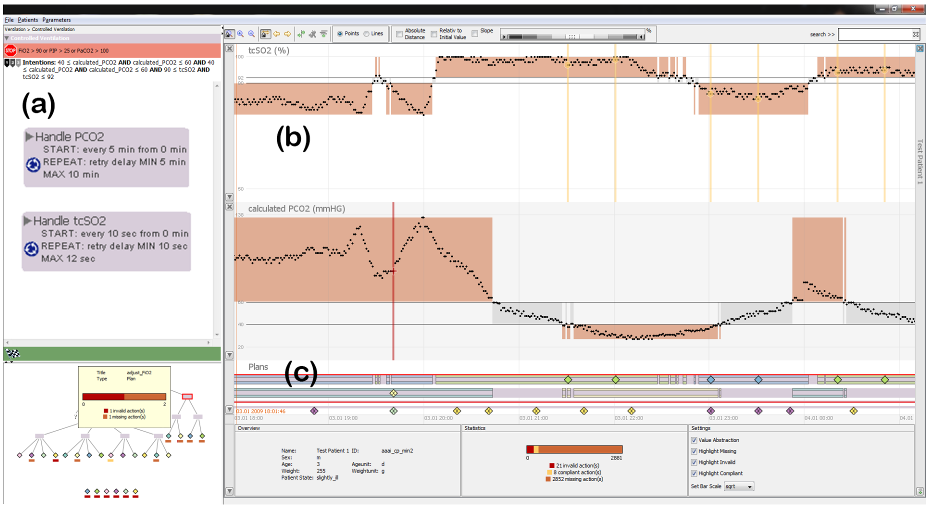

Fig. 1. Visual analysis of guideline conformance [8]. The recommended care process is visualized in a flow chart like graph in (a). In this screenshot the graph is zoomed in to show two repeated actions that should be carried out in parallel. (b) shows the patient's parameters that are affected by these actions. (c) represents actual clinical actions on a time axis. Actions that conform to the guideline are indicated by diamonds, non-conform actions are diamonds marked with an X, and the delay of actions is indicated by colored horizontal bars. A pop-up window gives the user details about these actions when hovering them.

2.2 Plan Strips

Another example from the medical domain is the Plan Strips visualization [39]. This interactive approach visualizes the complex hierarchies of processes – again using the example of clinical practice guidelines – by a set of nested rectangles. These guidelines recommend treatment plans (groups of clinical actions) which are composed of sub-plans. The execution order of a plan determines in which of the following ways its sub-plans are to be carried out: in sequence, in parallel (all sub-plans start at the same time), unordered (in sequence, in parallel, or a mixture of both), or in any-order (in sequence but the order is not defined) [40]. In addition, cyclical plans represent plans which are carried out in loops. Plan Strips encode these different kinds of sub-plan synchronization by color (see Figure 2): Gray rectangles indicate recommended clinical actions or activities while colored containers represent the corresponding parent-plans as well as the execution order. Interactive selection and highlighting techniques are used to ease navigation. One limitation of this approach is that the visualization demands for some learning effort, however, Plan Strips demonstrate an alternative way of representing highly specified and complex conditions of a process model in a space efficient way.

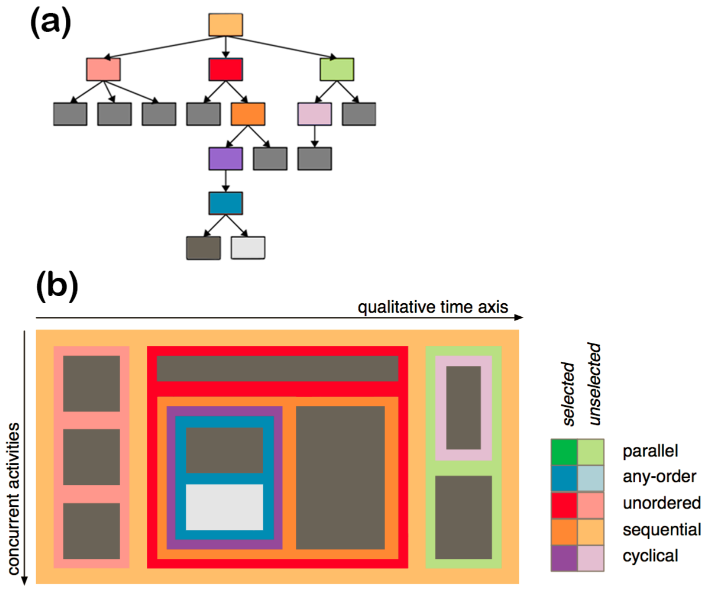

Fig. 2. Plan Strips [39] represent the hierarchical structure (a) of clinical plans and subplans as nested rectangles (b). User performed actions are indicated by gray rectangles (light gray when selected), while the color of parent containers indicate the execution order.

2.3 EventExplorer

EventExplorer [7] was designed to aid the analyst in exploring event log data, in order to gain some understanding of the data and the dependencies between events before deciding for concrete analysis methods or process mining algorithms. To this end, it combines pattern mining with interactive visualizations. Moreover, it supports the exploration of events on a temporal scale to aid the identification of important temporal dependencies of events (e.g., two events always occur together). Figure 3 shows a screenshot of EventExplorer. Each case (i.e., sequence of events) of the dataset is represented by a horizontal line composed of colored rectangles which represent the different events: The color indicates the type of event, while the order indicates the sequence of events.

The user can switch the representation to lay out events along a time axis to investigate their precise temporal behavior. EventExplorer uses pattern mining techniques to identify patterns within sequences of events. When a pattern is selected, all instances of this pattern within one case are highlighted and connected by gray arcs which emphasize the recurrence of this pattern within the case. EventExplorer also supports fuzzy pattern mining by allowing the use of wildcards when entering a pattern of interest. The approach presents interactive visualizations to (1) get familiar with the data, (2) to cluster and sort cases, and (3) to identify temporal patterns. The arcs visually emphasize recurring patterns as well as temporal distances (when using the time axis view).

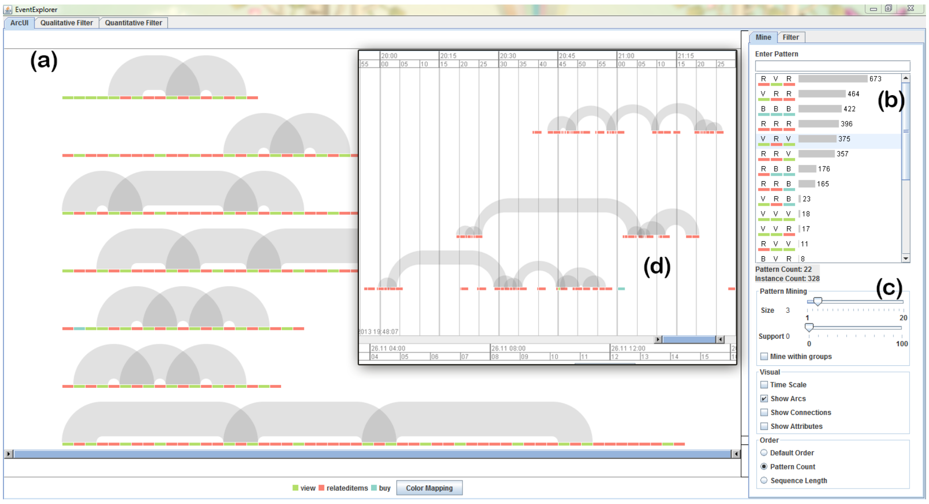

Fig. 3. EventExplorer [7]. In (a) EventExplorer represents all cases of the dataset as horizontal lines with color coded events (scrollable). (b) shows the different patterns (together with their number of occurrences) which match the settings in (c), i.e., the length of the pattern and the minimum number of occurrences. Alternatively, the user can enter a specific pattern of interest into the text field in (b). Selected patterns (in this case 'V-R-V') are highlighted by gray arcs in (a). The user can switch the x-axis in (a) from a sequential axis to a time axis (d) in order to explore the precise temporal characteristics of event sequences.

The three approaches outlined in detail were designed for very diverse purposes and give only examples of applications of visual analytics in process mining. However, there are still many open challenges but also opportunities when looking at the combination of these two research fields. We derived these challenges from the literature review, our previous work, and our long lasting experience in the research field of visual analytics.

3 Challenges and Opportunities

Researchers have identified challenges of today's process mining research in various works (e.g., [33, 47]). In [47] van der Aalst et al. even mention the potential of visual analytics to support process mining. Others again have summarized open challenges in the field of visual analytics (e.g., [30, 45, 50]). When focusing on the combination of visual analytics and process mining, we face related challenges.

C1: Intertwining Process Mining with Visual Analytics

Van der Aalst et al. identify the combination of process mining with other types of analysis as one of the challenges in process mining [47]. Moreover, they explicitly mention the potential of process mining combined with visual analytics for leading to further insights. Yet, how to intertwine process mining with visual analytics techniques is still an open problem. There are a couple of approaches which already tackle this problem. Among these visual approaches to aid process mining, the use of flow charts [36] or directed graphs to represent the process model are most common (e.g., [20, 32, 22, 21, 39]). Visualizing the process model helps to understand and fine-tune the model, but also to check conformance. Other visual approaches show an overview of individual event sequences (e.g., [43, 26, 51, 48, 7]) which can be used to identify common paths and thus, to derive a process model. Visualizations that show the conformance of individual event sequences (e.g., [8, 1]), however, help to identify shortcomings of the existing model, but also variations of individual cases. Yet, there is still much unexploited potential of visual analytics to better support pattern discovery and the derivation of process models, fine-tuning and enhancing these models, as well as identifying conformance problems and finding alternative solutions. To this end, visualizations need to be tightly coupled to analytical techniques and mining algorithms in an iterative loop, giving immediate visual feedback to changes and adjustments made to the algorithms. Appropriate views and interactions need to be tailored to specific user groups with respect to their specific tasks and data [35].

C2: Scalability and Aggregation

The scalability of visualizations is a well-known challenge in visual analytics [30, 45, 50]. This is also true in the context of process mining, which usually deals with huge amounts of event sequences that need to be analyzed. On the one hand, process mining algorithms tackle this problem by reducing these huge event logs to a manageable amount of patterns and process models. On the other hand, the analysis of the raw data and single event sequences may yield some important insights, and thus, visualizations need to support the simultaneous analysis of huge amounts of individual event sequences. There are some approaches that tackle this problem to some extent (e.g., [26, 51, 7]) by simultaneously showing a number of individual event sequences at minimum display space. Still the number of individual event sequences that can be displayed simultaneously is limited and it is an open challenge how to design interactive visualizations to support the whole analysis cycle from single event sequences to multiple event logs. There is a need for scalable visualizations able to represent a big number of cases while allowing zoom-in and drill-down to investigate details about single event sequences, in analogy to the famous Information Seeking Mantra by Shneiderman: *'overview first, zoom and filter, then details-on-demand'* [41, p.337].

C3: Interaction to Support Process Discovery and Enhancement

This challenge is tightly coupled with the last challenge of finding appropriate visualization techniques to support the different tasks. Data exploration is an iterative process with trial and error loops. Suitable interaction techniques need to support drilling down from an overview visualization to the investigation of single events. Commonly used techniques include zoom and filter functionalities, however, other interaction techniques may be needed depending on the data, user, and task [35]. Further interaction techniques include *select*, *explore*, *reconfigure*, *encode*, *abstract/elaborate*, and *connect* [52]. Allowing the analyst to interactively investigate the raw event data as well as the effects of changes made to models and mining algorithms may foster new insights and improve process mining results.

EventExplorer [7] allows for sorting cases by pattern count or by time. Another interesting way of sorting would be by similarity of event sequences in order to aid the identification of different groups of cases. Moreover, it supports the investigation of events in sequence or on a temporal scale. An interesting addition would be allowing the analyst to apply different ways of aligning cases vertically, for instance, aligning cases with the first appearance of a specific event.

C4: Data Quality and Uncertainty

Data quality is a general challenge in visual analytics [27], as well as in process mining [47], and thus, data quality is just as important when it comes to the combination of both fields. While process mining algorithms require good data quality and well-structured event logs, in practice these can be erroneous and badly structured. Instances of the same event type may have different names, event granularities may vary, and quite often the event log contains missing, incorrect, imprecise, uncertain, or irrelevant data. Data quality control can be divided into (1) data profiling (i.e., identifying and communicating quality problems), (2) data cleansing (i.e., correcting erroneous data), and (3) data transformation (i.e., transforming the data into appropriate formats for automatic processing) [19, 27]. There are visual approaches that tackle the problem of data quality (e.g., [17, 9, 29, 28]), however, the specific needs in process mining are hardly considered, such as case heterogeneity, event granularity, or concept drift.

On the other hand, data often contains some amount of uncertainty. Event log data may, for instance, contain uncertainties about which event type corresponds to which log entry or about the exact time of an event. Visually communicating these uncertainties to the analyst (e.g., [18]) is important for a better informed reasoning.

C5: Complexity of Time-Oriented Data

Usually event log data has an inherent temporal structure. For any such event log at least the sequence of events is known and used to analyze paths and processes. Yet, there are also event logs with precise timestamps for each event, allowing for a detailed analysis regarding the temporal dimension of events and event sequences. Time, however, is a complex data domain with very special features that require special consideration [2, 4, 41]. Some examples of these special characteristics are: time can be given as time points or intervals, conventionally time is aggregated irregularly (i.e., 60 minutes per hour, 24 hours per day, 28-31 days per month, 356/366 days per year, ...), leap seconds and days, different time zones, recurrences (e.g., seasonal cycles), and it has a strong social component (e.g., weekends, holidays) [2]. Ailenei et al. [3] outline use cases with special consideration of the temporal behavior of processes that need to be considered in process mining. When it comes to the visualization of time-oriented data, Aigner et al. [2] give a comprehensive overview. However, considering these peculiar characteristics of events (e.g., type of day, some cyclic behavior synchronized with the calendar, or events that co-occur with a certain delay) and adequately visualizing them for the task at hand is still an open challenge.

C6: Evaluation

Since visual analytics aims at supporting human reasoning and gaining new insights, the quality of visual analytics approaches is hard to quantify. Plaisant [37] outlines different challenges of evaluating visualizations and emphasizes that evaluation strategies need to take the exploratory nature of tasks and the added value of visualization such as overall awareness and potential discoveries into account. Moreover, praxis related aspects such as a successful adoption of these visualizations need to be considered. There are different approaches for evaluating visualizations [42, 25, 15] which usually involve target user participation and are highly task dependent (see [4] for a task framework and [31] for a time-specific task framework). In any case it is recommended to pursue an iterative design approach which involves feedback of the target user as early as possible.

On the other hand, evaluation in process mining is an open challenge too. There are different approaches on how to evaluate derived process models (e.g., [38, 6]), and Ailenei et al. [3] present a set of task specific use cases for the evaluation of process mining tools. These use cases may as well help to evaluate visual analytics solutions for process mining from a technical point of view. However, the evaluation of gained awareness and insights is still an open challenge.

4 Conclusion

Handling huge amounts of process data can be tackled in different ways: automatic mining techniques, visualizations, or by visual analytics which combines the benefits of both fields. Especially the latter has potential to improve the results of process analysis in many ways. In this paper I highlighted some examples of how this was approached in recent years. Moreover, I outlined six open challenges which reveal opportunities for further research. I believe, visual analytics is beneficial to many tasks of process mining and – if done right – will have substantial impact on the quality of process analysis.

ACKNOWLEDGMENTS

The research leading to these results has received funding from the Centre for Visual Analytics Science and Technology CVAST, funded by the Austrian Federal Ministry of Science, Research, and Economy in the exceptional Laura Bassi Centres of Excellence initiative (#822746).

References

1. A. Adriansyah, B. van Dongen, and W. M. P. van der Aalst. Conformance checking using cost-based fitness analysis. In *Proceedings of the Enterprise Computing Conference (EDOC) 2011*, pages 55–64, Piscataway, NJ, USA, 2011. IEEE Educational Activities Department.
2. W. Aigner, S. Miksch, H. Schuman, and C. Tominski. *Visualization of Time-Oriented Data.* Human-Computer Interaction. Springer Verlag, 1st edition, 2011.
3. I. Ailenei, A. Rozinat, A. Eckert, and W. M. P. van der Aalst. Definition and validation of process mining use cases. In F. Daniel, K. Barkaoui, and S. Dustdar, editors, *Business Process Management Workshops: BPM 2011 International Workshops, Revised Selected Papers, Part I*, pages 75–86, Berlin/Heidelberg, Germany, 2012. Springer.
4. N. Andrienko and G. Andrienko. *Exploratory Analysis of Spatial and Temporal Data: A Systematic Approach.* Springer, Berlin/Heidelberg, Germany, 2006.
5. F. J. Anscombe. Graphs in statistical analysis. *The American Statistician*, 27(1):17–21, 1973.
6. D. Ardagna, M. Mecella, and J. Yang, editors. *New Quality Metrics for Evaluating Process Models*, Berlin/Heidelberg, Germany, 2009. Springer.
7. P. Bodesinsky, B. Alsallakh, T. Gschwandtner, and S. Miksch. Exploration and assessment of event data. In E. Bertini and J. C. Roberts, editors, *Sixth International EuroVis Workshop on Visual Analytics (EuroVA) 2015*, page 5. The Eurographics Association, The Eurographics Association, 2015.
8. P. Bodesinsky, P. Federico, and S. Miksch. Visual analysis of compliance with clinical guidelines. In *Proceedings of the 13th International Conference on Knowledge Management and Knowledge Technologies (i-KNOW '13)*, pages 12:1–12:8. ACM, ACM, 2013.

9. C. Bors, T. Gschwandtner, S. Miksch, and J. Gärtner. Qualitytrails: Data quality provenance as a basis for sensemaking. In K. Xu, S. Attfield, and T. J. Jankun-Kelly, editors, *Proceedings of the IEEE VIS Workshop on Provenance for Sensemaking*, pages 1–2, 2014.
10. S. K. Card, J. D. Mackinlay, and B. Shneiderman, editors. *Readings in Information Visualization: Using Vision to Think*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 1999.
11. Conferences on Business Process Management. `http://bpm-conference.org/BpmConference/`. Accessed: 2016-08-25.
12. EuroVis (EG/VGTC Conference on Visualization). `https://www.eg.org/index.php/about-eg/working-groups/67-about-eg/working-groups/273-working-group-on-data-visualization-events`. Accessed: 2016-08-25.
13. D. A. Ferrucci. Introduction to "This is Watson". *IBM Journal of Research and Development*, 56(3.4):1–15, May 2012.
14. M. J. Field and K. N. Lohr, editors. *Clinical Practice Guidelines: Directions for a New Program*. National Academies Press, Institute of Medicine, Washington DC, 1990. `http://www.nap.edu/books/0309043468/html/` (last accessed: April, 2016).
15. C. Forsell and M. Cooper. *Handbook of Human Centric Visualization*, chapter An Introduction and Guide to Evaluation of Visualization Techniques Through User Studies, pages 285–313. Springer New York, New York, NY, 2014.
16. Google Scholar. `https://scholar.google.at/`. Accessed: 2016-08-25.
17. T. Gschwandtner, W. Aigner, S. Miksch, J. Gärtner, S. Kriglstein, M. Pohl, and N. Suchy. TimeCleanser: A visual analytics approach for data cleansing of time-oriented data. In S. Lindstaedt, M. Granitzer, and H. Sack, editors, *14th International Conference on Knowledge Technologies and Data-driven Business (i-KNOW 2014)*, pages 1–8. ACM Press, ACM Press, 2014.
18. T. Gschwandtner, M. Bögl, P. Federico, and S. Miksch. Visual encodings of temporal uncertainty: A comparative user study. *IEEE Transactions on Visualization and Computer Graphics*, 22:539–548, Jan. 2016.
19. T. Gschwandtner, J. Gärtner, W. Aigner, and S. Miksch. A taxonomy of dirty time-oriented data. In G. Quirchmayr, J. Basl, I. You, L. Xu, and E. Weippl, editors, *Lecture Notes in Computer Science (LNCS 7465): Multidisciplinary Research and Practice for Information Systems (Proceedings of the CD-ARES 2012)*, pages 58 – 72. Springer, Berlin / Heidelberg, Springer, Berlin / Heidelberg, 2012.
20. C. W. Günther and W. M. P. Aalst. Fuzzy mining – adaptive process simplification based on multi-perspective metrics. In G. Alonso, P. Dadam, and M. Rosemann, editors, *Proceedings of the 5th International Conference on Business Process Management (BPM 2007)*, pages 328–343, Berlin, Heidelberg, Sept. 2007. Springer Berlin Heidelberg.
21. C. W. Günther and A. Rozinat. Disco: Discover your processes. In *Proceedings of the Demonstration Track of the 10th International Conference on Business Process Management (BPM 2012), Tallinn, Estonia, September 4, 2012*, pages 40–44, 2012.
22. M. Hipp, B. Mutschler, and M. Reichert. Navigating in process model collections: A new approach inspired by google earth. In F. Daniel, K. Barkaoui, and S. Dustdar, editors, *Business Process Management Workshops*, volume 100 of *Lecture Notes in Business Information Processing*, pages 87–98. Springer Berlin Heidelberg, 2012.
23. IEEE VIS. `http://ieeevis.org/`. Accessed: 2016-08-25.
24. IEEE Xplore Digital Library. `http://ieeexplore.ieee.org/Xplore/home.jsp`. Accessed: 2016-08-25.

25. T. Isenberg, P. Isenberg, J. Chen, M. Sedlmair, and T. Möller. A systematic review on the practice of evaluating visualization. *IEEE Transactions on Visualization and Computer Graphics*, 19(12):2818–2827, Dec. 2013.
26. R. P. Jagadeesh Chandra Bose and W. M. P. van der Aalst. Process diagnostics using trace alignment: Opportunities, issues, and challenges. *Inf. Syst.*, 37(2):117–141, Apr. 2012.
27. S. Kandel, J. Heer, C. Plaisant, J. Kennedy, F. van Ham, N. H. Riche, C. Weaver, B. Lee, D. Brodbeck, and P. Buono. Research directions in data wrangling: Visuatizations and transformations for usable and credible data. *Information Visualization*, 10(4):271–288, Oct. 2011.
28. S. Kandel, A. Paepcke, J. Hellerstein, and J. Heer. Wrangler: Interactive visual specification of data transformation scripts. In *Proc. of the ACM Conference Human Factors in Computing Systems (CHI 2011)*, pages 3363–3372, May 2011.
29. S. Kandel, R. Parikh, A. Paepcke, J. Hellerstein, and J. Heer. Profiler: Integrated statistical analysis and visualization for data quality assessment. In *Proc. of the International Working Conference on Advanced Visual Interfaces (AVI'12)*, pages 547–554, May 2012.
30. D. Keim, G. Andrienko, J. Fekete, C. Grg, J. Kohlhammer, and G. Melanon. Visual Analytics: Definition, process, and challenges. In A. Kerren, J. T. Stasko, J. Fekete, and C. North, editors, *Information Visualization: Human-Centered Issues and Perspectives*, LNCS 4950, pages 154–175. Springer, Berlin/Heidelberg, Germany, 2008.
31. T. Lammarsch, A. Rind, W. Aigner, and S. Miksch. Developing an extended task framework for exploratory data analysis along the structure of time. In K. Matkovic and G. Santucci, editors, *Proceedings of the EuroVis Workshop on Visual Analytics in Vienna, Austria (EuroVA 2012)*, pages 31–35. Eurographics, Eurographics, 2012.
32. E. Maguire, P. Rocca-Serra, S. Sansone, J. Davies, and M. Chen. Visual compression of workflow visualizations with automated detection of macro motifs. *IEEE Transactions on Visualization and Computer Graphics*, 19(12):2576–2585, 2013.
33. R. S. Mans, W. M. P. Aalst, R. J. B. Vanwersch, and A. J. Moleman. Process mining in healthcare: Data challenges when answering frequently posed questions. In R. Lenz, S. Miksch, M. Peleg, M. Reichert, D. Riaño, and A. Teije, editors, *Process Support and Knowledge Representation in Health Care: BPM 2012 Joint Workshop (ProHealth 2012/KR4HC 2012)*, pages 140–153, Berlin, Heidelberg, Sept. 2013. Springer Berlin Heidelberg.
34. R. Mazza. Perception. In *Introduction to Information Visualization*, chapter 3, pages 33–44. Springer-Verlag, Limited, 2009.
35. S. Miksch and W. Aigner. A matter of time: Applying a data-users-tasks design triangle to visual analytics of time-oriented data. *Computers & Graphics, Special Section on Visual Analytics*, 38:286–290, 2014.
36. K. Peters. http://home.southernct.edu/p̃etersk1/csc400/csc400-flowchart.htm. (Accessed: 2009-01-05).
37. C. Plaisant. The challenge of information visualization evaluation. In *Proceedings of the Working Conference on Advanced Visual Interfaces (AVI '04)*, pages 109–116, New York, NY, USA, 2004. ACM.
38. A. Rozinat, A. K. A. de Medeiros, C. W. Günther, A. J. M. M. Weijters, and W. M. P. van der Aalst. The need for a process mining evaluation framework in research and practice. In A. ter Hofstede, B. Benatallah, and H.-Y. Paik, editors, *Business Process Management Workshops: BPM 2007 International Workshops, Revised Selected Papers*, pages 84–89, Berlin/Heidelberg, Germany, 2008. Springer.

39. A. Seyfang, K. Kaiser, T. Gschwandtner, and S. Miksch. Visualizing complex process hierarchies during the modeling process. In R. Brown, S. Kriglstein, and S. Rinderle-Ma, editors, *First International Workshop on Theory and Applications of Process Visualization (TAProViz'12)*, Lecture Notes in Business Information Processing (LNBIP). Springer, Springer, 2012.
40. Y. Shahar, S. Miksch, and P. Johnson. The Asgaard Project: A task-specific framework for the application and critiquing of time-oriented clinical guidelines. *Artificial Intelligence in Medicine*, 14:29–51, 1998.
41. B. Shneiderman. The eyes have it: A task by data type taxonomy for information visualizations. In *Proceedings of the 1996 IEEE Symposium on Visual Languages*, pages 336–343, Piscataway, NJ, USA, September 1996. IEEE Educational Activities Department.
42. B. Shneiderman and C. Plaisant. Strategies for evaluating information visualization tools: Multi-dimensional in-depth long-term case studies. In *Proceedings of the 2006 AVI Workshop on BEyond Time and Errors: Novel Evaluation Methods for Information Visualization*, BELIV '06, pages 1–7, New York, NY, USA, 2006. ACM.
43. M. Song and W. van der Aalst. Supporting process mining by showing events at a glance. In *Proceedings of the 17th Annual Workshop on Information Technologies and Systems (WITS)*, pages 139–145, 2007.
44. I. Spence and R. F. Garrison. A remarkable scatterplot. *The American Statistician*, pages 12–19, 1993.
45. J. J. Thomas and K. A. Cook, editors. *Illuminating the Path: The Research and Development Agenda for Visual Analytics.* IEEE Educational Activities Department, Piscataway, NJ, USA, 2005.
46. W. van der Aalst, editor. *Process Mining: Discovery, Conformance and Enhancement of Business Processes.* Springer, Berlin/Heidelberg, Germany, 1st edition edition, 2011.
47. W. M. P. van der Aalst et al. Process mining manifesto. In F. Daniel, K. Barkaoui, and S. Dustdar, editors, *Business Process Management Workshops*, volume 99 of *Lecture Notes in Business Information Processing*, pages 169–194. Springer Berlin Heidelberg, 2012.
48. K. Vrotsou, J. Johansson, and M. Cooper. Activitree: Interactive visual exploration of sequences in event-based data using graph similarity. *IEEE Transactions on Visualization and Computer Graphics*, 15(6):945–952, Nov. 2009.
49. Wikipedia. http://en.wikipedia.org/wiki/gantt_chart. (Accessed: 20011-12-22).
50. P. C. Wong, H. Shen, C. R. Johnson, C. Chen, and R. B. Ross. The top 10 challenges in extreme-scale visual analytics. *IEEE Computer Graphics and Applications*, 32(4):63–67, 2012.
51. K. Wongsuphasawat and D. Gotz. Exploring flow, factors, and outcomes of temporal event sequences with the outflow visualization. *IEEE Transactions on Visualization and Computer Graphics*, 18(12):2659–2668, 2012.
52. J. S. Yi, Y. ah Kang, J. T. Stasko, and J. A. Jacko. Toward a deeper understanding of the role of interaction in information visualization. *IEEE Transactions on Visualization and Computer Graphics*, 13(6):1224–1231, 2007.