Eurographics Conference on Visualization (EuroVis) 2017
B. Kozlíková, T. Schreck, and T. Wischgoll (Editors)

Short Paper

On Establishing Visualization Requirements: A Case Study in Product Costing

Z. Vosough[1,2], R. Groh[2] and H.-J. Schulz[3]

[1]SAP ICN Research & Innovation Dresden, Germany
[2]TU Dresden, Germany
[3]University of Rostock, Germany

Abstract

The process of identifying visualization requirements is an important part of every visualization researcher's and practitioner's job. Nevertheless, the scientific literature is rather sparse on this topic, usually resorting to some form of user-centered design that is rarely further detailed. In this paper, we give an account of our procedure, our results, our problems and solutions for gathering visualization requirements in an ongoing business project to introduce visualization to the field of product costing. By providing insight in our experiences and extracting general points of advice from them, we aim to give some practical guidance for establishing requirements in real-world visualization projects.

Categories and Subject Descriptors (according to ACM CCS): I.3.3 [Computer Graphics]: Picture/Image Generation—Line and curve generation

1. Introduction

There is a sizable body of research on how to use visualization in *requirements engineering* – i.e., the process of identifying, specifying, and verifying requirements for a system under construction [GMM07, CLGG09, RRB*14]. Yet there are only a few literature sources dedicated to requirements engineering in visualization. For all practical purposes, these sources usually assume a *user-centered visualization design* process [TM04, KKUW06, KSDK11, RHR16]. Its result is produced through an iterative negotiation between the needs of the end users and the technical and representational possibilities available to the visualization designer. For this process to succeed, the literature stresses the importance of having direct access to end users who are willing to invest their time in the design process, as well as having access to the actual or at least realistic data [SMM12, SD12]. This may actually be attainable in academia, where visualization researchers work jointly and openly with researchers from other domains as their users. But in industrial collaborations, visualization design gets "messy, iterative, and complex" [MMAM14], as the visualization researcher has to deal with managerial hierarchies and issues of confidentiality in addition to the engineering challenges [SIBB11, CGM16].

In this paper, we report on these added "complexities" as we encountered them during our ongoing development of a visualization solution for SAP product lifecycle costing (PLC). In a nutshell, *product lifecycle costing* subsumes all methods for estimating and optimizing the costs a product incurs over its lifetime – from its initial design, to manufacturing, using, maintaining, and retiring it. The idea behind this procedure is to reduce costs as early as possible, as the bulk of a product's total cost is determined by early design decisions, such as which parts and materials to use [AG98].

The contribution of this paper is threefold:

- It introduces the case study and retraces the steps we have taken to establish visualization requirements from the initial idea to the first prototypes in Section 2.
- It reports on the problems we encountered during these steps and how we resolved, circumvented, or mitigated them in Section 3.
- It recollects some lessons learned from this case study that may be helpful to others in Section 4.

2. Visualization Requirements for Product Costing

This section gives a streamlined account of our process of requirements elicitation from more than 30 co-innovation customers who give regular input on SAP PLC product ideas and prototypes. This process of "co-innovation" allows for discussing experimental features and getting early feedback in a user and task-based design approach [TM04]. As approximately every quarter a new software version is released, we run regular co-innovation workshops with customers in the same quarterly rhythm. The visualization requirements were established in four steps: current visualization practice, visualization tasks, characteristics of the visualized data, and further detailed requirements by means of discussing first prototypes. The following subsections will briefly outline our procedure and results for each of them.

DOI: 10.2312/eurovisshort.20171140

2.1. Current Practice

Procedure: Our first step was to understand our customers' current solutions and needs with respect to visualization support. Note that this step is often left out, with most projects going right into task and data requirements. We decided for this extra step upfront to make sure that our research is not based on a false assumption of a need for visualization where there actually is none [Mun09]. To that end, we conducted a survey with 21 co-innovation customers. They were asked broadly which product costing solutions they currently employ, which data visualizations they use, as well as for which parts of the costing analysis process they use data visualization. The survey results were complemented with personal interviews conducted with customers at co-innovation workshops.

Results: The collected answers confirmed our subjective experiences that most customers currently use spreadsheet solutions and the common visualization techniques offered by them – predominantly pie charts and bar charts – for costing analyses. Yet, the charts are utilized almost exclusively at the end of the analysis process for reporting and presentation purposes, with the only exception being visual comparisons that are occasionally performed during the analysis. While spreadsheets and their associated visualizations match the needs of analyzing tabular data, cost analyses involve much more than that. For example, choosing a certain material for a product may mean that it is cheaper to buy, but more delicate to handle and thus produce more defective goods during manufacturing. Tracing such dependencies and observing the repercussions of different design choices can hardly be captured in a tabular form. Hence, we saw indeed an opening for a tailored visualization.

2.2. Visualization Tasks

Procedure: To substantiate our hypothesis that parts of the cost analysis are ill-suited to be pursued with a tabular data display alone, we conducted hour-long group discussions with a total of 30 participants from 16 companies during co-innovation workshops in Germany and in the US. The discussions focused on the end users' tasks that are perceived as complex, that take a long time to perform, or that are currently simply impossible to carry out, as these tasks could potentially benefit from visualization support. Note that while the scientific literature abounds with abstract visualization tasks and taxonomies, there is little research on how to conduct a visualization task analysis. We loosely followed the general recommendations given in [KKUW06].

Results: From these discussions, four main tasks emerged:

T1: To identify the main cost drivers by comparing multiple cost calculations with each other, so as to gauge the impact of adding or removing individual items or assemblies on the overall costs.
T2: To see how close the calculation is to a defined cost target and which assemblies are above or below targets – including prognostics, i.e., which assemblies get less or more costly over time.
T3: To determine incomplete or inconsistent cost calculations – e.g., to find missing prices and see where prices have been estimated instead of being derived from reliable master data.
T4: To assess the reliability of the overall cost calculated from price sources with different confidence levels. Determine best, realistic, and worst cases based on projections for future costs.

2.3. Data Structure and Size

Procedure: To better understand the datasets on which our customers operate, a survey was conducted with 12 customers. Again, while there exists a number of different papers on metadata, dataspace notations, and data descriptors for visualization, literature on the best practices of collecting this information from users is hard to find. To shape our survey, we took inspiration from two studies conducted previously on enterprise data analysis [KPHH12, KBHP14].

Results: It turns out that the costing data structure is hierarchical and additive, as the overall product cost is basically the sum of the costs of its parts, and so on for sub-parts and finally raw materials and labor. Furthermore, the product cost consists of several calculations, with each calculation existing in several versions. These calculation versions (CV) are used to take different scenarios into account, such as optimistic vs. pessimistic price dynamics, so that the cost development of a product can be projected into the future and factored into the analysis. As it can be seen from Table 1, the range among the different customers with respect to their data sizes is rather large. The reason is, that some companies build new products based on previous product versions and thus they can analyze their product costs by simply adding a slightly changed calculation version for each new product. Whereas other companies build completely different products every time and thus have to add entirely new costing structures with new calculations for each.

	Min	Typical	Max
No. calculations	10	100	300K
No. CV per calculation	1	5	120
Size of a CV	20	1K	900K
Tree depth	3	5	20
No. materials	40K	400K	3Mio

Table 1: *Data sizes as surveyed from 12 different customers.*

2.4. Early Visualization Prototypes

Procedure: To deepen the discussion with the customers, we needed to produce something tangible that allowed to break free from the mental map of bar charts and pie charts that they had in mind. Building early prototypes [SD12, KSDK11] and mock-ups [RHR16] are common strategies in visualization design, where end users from application domains often decide on visualizations in an "I-know-it-when-I-see-it" manner. Hence, we developed two prototypes shown in Figure 1 for substantiating the discussion of task, data, and representation: a Squarified Treemap [BHvW00] and a Sankey diagram [RHF05]. The Treemap is an obvious choice for attribute-centric tasks, such as looking at cost drivers [T1] or target costs [T2], while still keeping the costs hierarchically organized through the layout. As Treemaps are known to scale quite well [FP02], they are a good fit for the wide range of tree sizes specified by the customers. The Sankey diagram is more geared towards structure-centric tasks, such as tracing inconsistencies to their origin [T3] or analyzing the propagation of reliability scores through the hierarchical cost structure [T4], as it provides a clear view of the costing structure. Both visualization prototypes were outfitted with basic interactions, such as folding/unfolding hierarchical branches and tooltips for detailed information. We discussed both prototypes in an informal setting at two customer workshops.

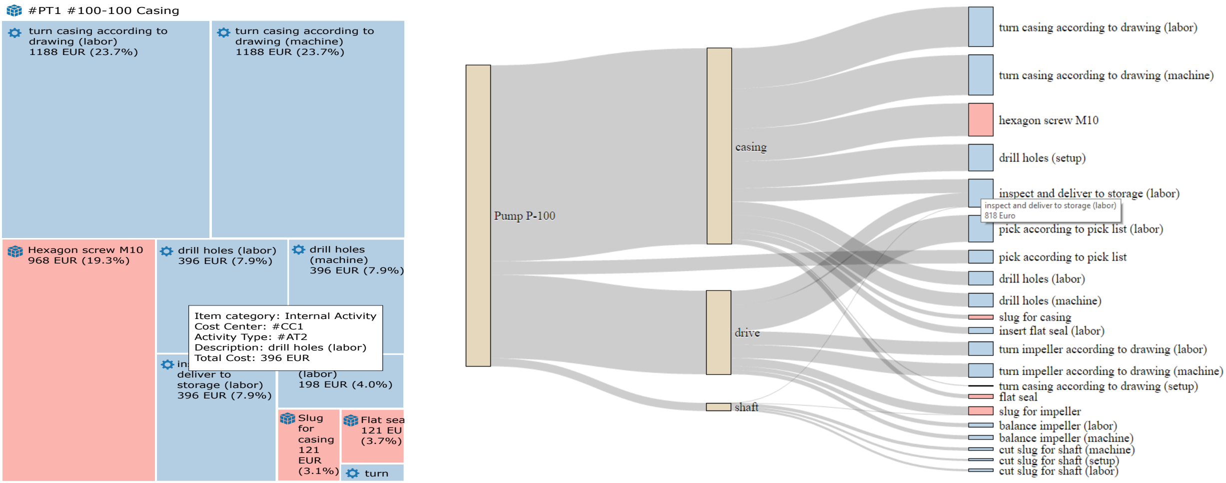

Figure 1: *Early visualization prototypes – (left) a Treemap and (right) a Sankey diagram. Both figures show a small dataset of an industrial pump. The blue components represent activities, whereas the pink components indicate materials.*

Results: The most interesting aspect that surfaced during these discussions was a preference of the customers towards the Sankey diagram. One reason for this tendency was that while the costing structure is hierarchical in nature, it still occurs that items of the same type – oftens screws or bolts – are used in the assembly of different components of the product. Since the Treemap is strictly hierarchical, these parts are shown individually for each component in multiple places, which makes it hard to judge their overall impact on the product cost. Yet the Sankey diagram allows for a more flexible representation and it can be seen in Figure 1 that the item "inspect and deliver to storage" is connected to different extent to all three of the higher level components "casing", "drive", and "shaft". The discussion with the customers made us aware of how important it is to the users to see them in such a combined way.

3. Challenges in Yielding Visualization Requirements

The account of our experiences given in the previous section is by far not complete. As it is commonly practice in academic literature, it tells the "success story", but leaves out the challenges we ran into. In particular as the literature on establishing visualization requirements is sparse, we deem it important to report on those challenges and how we solved them. Papers on visualization design already mention some general pitfalls to avoid and give a few recommendations to follow. If we had followed them, we would have terminated the project at an early stage:

- Many co-innovation customers turned out not to be end users, but "gatekeepers" [SMM12] who at best fit the roles of "human-as-viewer" [WLC16] or "consumers of analysis" [KPHH12].
- Many of the customers were hesitant to share information for confidentiality reasons.
- Those who were willing to share were hard to reach and often unresponsive, as among their management duties, getting new visualization options was not high on the priority list.
- And from those few who were excited to help, we got mostly a wild mix of overly specific feature requests that changed with every new project they worked on.

As the co-innovation customers were the only ones we had access to, it was also not an option to find better collaboration partners, as suggested in some literature. Neither was terminating the project, as that would have meant to discard a promising research direction and a possible unique selling point for our costing software. So, we had to find creative ways to nevertheless get the necessary requirements for designing our visualization.

3.1. Getting Information from End Users

Having only limited access to actual day-to-day users makes a user-centered visualization design impossible. The literature lists alternatives ranging from *activity-centered design* to *goal-directed design* [Wil09]. Yet again, to get information about activities or goals, we would need end users to work with us.

In cases where we could not ask the end users themselves, we tried to ask the people who work with them to get "second-hand requirements". At SAP, this is the solutions management team who interacts with the companies in all cases that require technical support. They know the problems of the users first hand and could give us information ranging from the companies' software landscape to frequently asked questions. In addition, we sent out surveys to our co-innovation customers that were intentionally phrased in more technical jargon and asking for completion in a rather short timeframe. Putting up these additional hurdles ensured that only end users who know their costing software and analysis practices could easily fill out these surveys within the allotted time. Thus, we could be confident to only receive responses either from knowledgeable managers, or from costing analysts to whom the questionnaire was passed down. While such measures may increase the quality of responses, they unfortunately also decrease their quantity and thus have to be employed with care.

3.2. Overcoming the Hurdles of Information Sharing

The hesitance to share information is a wide-spread problem in requirements engineering for which there exists no simple solution [Gog93, GL93]. The field of Social Engineering has developed methods to build trust and convince people to disclose information [Had10]. Yet, many of these methods are deceptive bordering on the unethical, and stand in stark contrast to the good practices of open and honest customer communication.

To lower these hurdles, we followed a combination of the *employee-pull* and *researcher-push* strategies [SIBB11]. For the employee-pull part, we utilized the *toolsmith method* [vW06] making sure that our discussion focused on the needs of the customers and not on our needs for their data and feedback. As of course the needs of the customers can only be sensibly discussed in the context of their specific data and tasks, we got the needed information almost as a byproduct of the customer interviews. Once we had collected enough information to start producing early visualization prototypes, we switched to the researcher-push part, following the *design-first* method [PRN15]. This means that without yet having detailed knowledge of data or tasks, we already put out two prototypes showing synthetic data. These allowed the customers to move from discussing their confidential data structures and business questions to discussing our prototypes.

3.3. Collecting Information from Unresponsive Partners

We made the observation that, compared to other business-related surveys and cooperation requests, the co-innovation customers did not respond as eagerly to our surveys. For example, from the 21 customers surveyed in the beginning of the process, only 5 initially returned the filled-out questionnaires. This is problematic, as so few returned questionnaires not only raise doubts about their representativeness, but also introduce a high potential for non-response bias in the results [Dem90, p.66]. As many co-innovation customers are in middle managerial positions, it is no wonder that answering a survey on business-related questions comes more easily to them than one about the implications of product costing on visualization.

In practice, we used two modes of collecting the answers we needed: the email survey and personal interviews to follow up. The survey via email has the advantage of being asynchronous, so that the customers can complete it at their leisure. It also offers the possibility to speak with the actual analyst or even to delegate filling out the survey directly to them. The downsides are its non-committal nature, as well as the possibility that questions are misunderstood or not answered with the necessary detail. Personal interviews at customer workshops have just the opposite characteristics, which makes them perfect to complement the email surveys. Interviews were particularly well received for issues that would have required lengthy written explanations, but that could simply be demonstrated by pointing at the screen during an interview.

3.4. Yielding Focused and Relevant Information

It is no secret that customers have different goals than visualization designers: Where customers prioritize the needs of their specific current project, the visualization designers try to realize a visualization that supports costing analysis in general. Note that this gap between the special-purpose tool the customer requires and the general-purpose tool the researcher desires is different to the customer/researcher gaps observed by v.Wijk [vW06].

To find common ground and avoid accidental overfitting, we planned multiple quarterly discussions and interviews over the course of the project. Looking at which problems and questions reappear over the course of these multiple meetings helped us to separate principal requirements from the specific "problems of the day". These principal requirements served as a baseline describing the typical costing analysis, which usually means to find out what makes the product expensive (task T1) on datasets of typical size (Table 1). None of these commonplace requirements were particularly "exciting" – neither to the customers, nor to us. But it is this common denominator that a visualization should first and foremost be able to handle, while special cases can be added onto this "base visualization" in subsequent design iterations.

4. Lessons Learned and Outlook

Apart from the concrete approaches outlined above, we made the following observations that helped us to yield requirements from our customers and that may serve as general recommendations:

- **Interview participants in their native language.** We observed that participants are more talkative when speaking in their mother tongue. That was the reason why we set up two independent workshops in Germany and the US, so that each group of customers had the opportunity to partake in the group discussions and interviews in their native language.
- **Ask about problems, not solutions.** It is easier to get people to open up by asking them about their gripes and grievances with the current system, than having them suggest features they might want to use in the future. This is in line with existing findings that middle-level managers are more likely to perceive problems than to come up with potentially useful features [Sut02, p.89].
- **Provide a neutral context for discussions.** By presenting our early and imperfect visualization prototypes to the participants, we not only gave them something to criticize (see previous point on problems vs. solutions), but also a neutral context in which to frame their statements. This catalyzed their conversation with us, but also their discussion among each other, as they did no longer have to worry about disclosing internal company details.

There are many more challenges in requirements engineering that still need research and further reports from visualization practitioners on how to resolve them – for example, how to manage conflicting requirements? While some of our observations and lessons learned seem to be tacit knowledge in the visualization community, we deem it important to make them explicit and we encourage other authors to do so as well. In this way, these guidelines can be collected and properly debated in meta studies and surveys to someday yield the still missing compilation of best practices for working with visualization end users.

5. Acknowledgments

We thank all members and customers of SAP Product Lifecycle Costing for their input on different parts of this research. Special thanks go to Jochen Rode and Stefan Hesse for contributing their experiences and thoughts, Marc Kirchhoff for conducting the survey on data sizes, Eliza Koshtoyan for implementing the Treemap prototype, and Martin Luboschik and the anonymous reviewers for their constructive feedback on the paper.

References

[AG98] ASIEDU Y., GU P.: Product life cycle cost analysis: State of the art review. *International Journal of Production Research 36*, 4 (1998), 883–908. doi:10.1080/002075498193444. 1

[BHvW00] BRULS M., HUIZING K., VAN WIJK J. J.: Squarified treemaps. In *Proceedings of the Joint Eurographics / IEEE TCVG Symposium on Visualization* (2000), de Leeuw W., van Liere R., (Eds.), Eurographics Association, pp. 33–42. doi:10.2312/VisSym/VisSym00/033-042. 2

[CGM16] CRISAN A., GARDY J. L., MUNZNER T.: On regulatory and organizational constraints in visualization design and evaluation. In *Proceedings of the BELIV Workshop: Beyond Time and Errors – Novel Evaluation Methods for Visualization* (2016), Sedlmair M., Isenberg P., Isenberg T., Mahyar N., Lam H., (Eds.), ACM, pp. 1–9. doi:10.1145/2993901.2993911. 1

[CLGG09] COOPER JR. J. R., LEE S.-W., GANDHI R. A., GOTEL O.: Requirements engineering visualization: A survey on the state-of-the-art. In *Proceedings of the 4th International Workshop on Requirements Engineering Visualization* (2009), Berenbach B., Gotel O., (Eds.), IEEE, pp. 46–55. doi:10.1109/REV.2009.4. 1

[Dem90] DEMING W. E.: *Sample Design in Business Research*. John Wiley & Sons, 1990. 4

[FP02] FEKETE J.-D., PLAISANT C.: Interactive information visualization of a million items. In *Proceedings of the IEEE Symposium on Information Visualization* (2002), Wong P. C., Andrews K., (Eds.), IEEE, pp. 117–124. doi:10.1109/INFVIS.2002.1173156. 2

[GL93] GOGUEN J. A., LINDE C.: Techniques for requirements elicitation. In *Proceedings of the IEEE International Symposium on Requirements Engineering* (1993), IEEE, pp. 152–164. doi:10.1109/ISRE.1993.324822. 4

[GMM07] GOTEL O. C. Z., MARCHESE F. T., MORRIS S. J.: On requirements visualization. In *Proceedings of the 2nd International Workshop on Requirements Engineering Visualization* (2007), Berenbach B., Gotel O., (Eds.), IEEE. doi:10.1109/REV.2007.4. 1

[Gog93] GOGUEN J. A.: Social issues in requirements engineering. In *Proceedings of the IEEE International Symposium on Requirements Engineering* (1993), IEEE, pp. 194–195. doi:10.1109/ISRE.1993.324858. 4

[Had10] HADNAGY C.: *Social Engineering: The Art of Human Hacking*. John Wiley & Sons, 2010. 4

[KBHP14] KANDOGAN E., BALAKRISHNAN A., HABER E. M., PIERCE J. S.: From data to insight: Work practices of analysts in the enterprise. *IEEE Computer Graphics and Applications 34*, 5 (2014), 42–50. doi:10.1109/MCG.2014.62. 2

[KKUW06] KULYK O., KOSARA R., URQUIZA J., WASSINK I.: Human-centered aspects. In *Human-Centered Visualization Environments*, Kerren A., Ebert A., Meyer J., (Eds.). Springer, 2006, pp. 13–75. doi:10.1007/978-3-540-71949-6_2. 1, 2

[KPHH12] KANDEL S., PAEPCKE A., HELLERSTEIN J. M., HEER J.: Enterprise data analysis and visualization: An interview study. *IEEE Transactions on Visualization and Computer Graphics 18*, 12 (2012), 2917–2926. doi:10.1109/TVCG.2012.219. 2, 3

[KSDK11] KOH L. C., SLINGSBY A., DYKES J., KAM T. S.: Developing and applying a user-centered model for the design and implementation of information visualization tools. In *Proceedings of the International Conference on Information Visualisation (IV)* (2011), Banissi E., Bertschi S., Burkhard R., Cvek U., Eppler M., Forsell C., Grinstein G., Johansson J., Kenderdine S., Marchese F. T., Maple C., Trutschl M., Sarfraz M., Stuart L., Ursyn A., Wyeld T. G., (Eds.), IEEE, pp. 90–95. doi:10.1109/IV.2011.32. 1, 2

[MMAM14] MCKENNA S., MAZUR D., AGUTTER J., MEYER M.: Design activity framework for visualization design. *IEEE Transactions on Visualization and Computer Graphics 20*, 12 (2014), 2191–2200. doi:10.1109/TVCG.2014.2346331. 1

[Mun09] MUNZNER T.: A nested model for visualization design and validation. *IEEE Transactions on Visualization and Computer Graphics 15*, 6 (2009), 921–928. doi:10.1109/TVCG.2009.111. 2

[PRN15] PAUL C. L., ROHRER R., NEBESH B.: A "design first" approach to visualization innovation. *IEEE Computer Graphics and Applications 35*, 1 (2015), 12–18. doi:10.1109/MCG.2015.7. 4

[RHF05] RIEHMANN P., HANFLER M., FROEHLICH B.: Interactive Sankey diagrams. In *Proceedings of the IEEE Symposium on Information Visualization* (2005), Stasko J., Ward M. O., (Eds.), IEEE, pp. 233–240. doi:10.1109/INFVIS.2005.1532152. 2

[RHR16] ROBERTS J. C., HEADLEAND C., RITSOS P. D.: Sketching designs using the five design-sheet methodology. *IEEE Transactions on Visualization and Computer Graphics 22*, 1 (2016), 419–428. doi:10.1109/TVCG.2015.2467271. 1, 2

[RRB*14] REDDIVARI S., RAD S., BHOWMIK T., CAIN N., NIU N.: Visual requirements analytics: a framework and case study. *Requirements Engineering 3*, 19 (2014), 257–279. doi:10.1007/s00766-013-0194-3. 1

[SD12] SLINGSBY A., DYKES J.: Experiences in involving analysts in visualisation design. In *Proceedings of the BELIV Workshop: Beyond Time and Errors – Novel Evaluation Methods for Visualization* (2012), Bertini E., Perer A., Lam H., Isenberg P., Isenberg T., (Eds.), ACM. doi:10.1145/2442576.2442577. 1, 2

[SIBB11] SEDLMAIR M., ISENBERG P., BAUR D., BUTZ A.: Information visualization evaluation in large companies: Challenges, experiences and recommendations. *Information Visualization 10*, 3 (2011), 248–266. doi:10.1177/1473871611413099. 1, 4

[SMM12] SEDLMAIR M., MEYER M., MUNZNER T.: Design study methodology: Reflections from the trenches and the stacks. *IEEE Transactions on Visualization and Computer Graphics 18*, 12 (2012), 2431–2440. doi:10.1109/TVCG.2012.213. 1, 3

[Sut02] SUTCLIFFE A.: *User-Centred Requirements Engineering*. Springer, 2002. doi:10.1007/978-1-4471-0217-5. 4

[TM04] TORY M., MÖLLER T.: Human factors in visualization research. *IEEE Transactions on Visualization and Computer Graphics 10*, 1 (2004), 72–84. doi:10.1109/TVCG.2004.1260759. 1

[vW06] VAN WIJK J. J.: Bridging the gaps. *IEEE Computer Graphics and Applications 26*, 6 (2006), 6–9. doi:10.1109/MCG.2006.120. 4

[Wil09] WILLIAMS A.: User-centered design, activity-centered design, and goal-directed design: A review of three methods for designing web applications. In *Proceedings of the ACM International Conference on Design of Communication (SIGDOC)* (2009), Mehlenbacher B., Protopsaltis A., Williams A., Slattery S., (Eds.), ACM, pp. 1–8. doi:10.1145/1621995.1621997. 3

[WLC16] WINTERS K. M., LACH D., CUSHING J. B.: A conceptual model for characterizing the problem domain. *Information Visualization 15*, 6 (2016), 301–311. doi:10.1177/1473871615608902. 3