Designing visualization software for super-wicked problems

Kirsten M. Winters\textsuperscript{a,*}, Judith B. Cushing\textsuperscript{b} and Denise Lach\textsuperscript{a}
\textsuperscript{a}Oregon State University, Corvallis, OR, USA
\textsuperscript{b}The Evergreen State College, Olympia, WA, USA

Abstract. Designing effective visualization software in the context of super wicked problems includes incorporating understanding of public policy and decision making settings into the software development process. Case study findings presented here show that a problem-driven design approach must include both the primary and secondary users of visualization software. The primary lesson learned is that extending the scope of the problem domain beyond the explicit functionality of creating visualizations, to include the reactions or enhanced participation of decision makers, will likely provide scientists with more effective software.

Keywords: Scientific visualization, visual analytics, wicked problems, climate change, post-normal science, software Development

1. Introduction

In order to explore how ecologists could use visualization to understand and communicate findings, social scientists partnered with a development team building software to suit these real users trying to solve real problems. Over the course of the development cycle, a qualitative case study was conducted to support lead software developers in understanding how and why primary users of software (the ecologists) intended to use visualizations of their data. Lead software developers originally had assumed that environmental scientists primarily used visualizations to enhance their own scientific inquiry, sorting out reams of sensed and modeled data to arrive at new insights into the underlying physical phenomena, or at least to conduct their own research more efficiently. Lead software developers also believed that scientists would use visualization to present results to other scientists. Within a year of the project’s inception, however, it became clear that two of the three environmental scientists, those studying the impacts of climate change, were using visualizations to explain complex environmental science phenomena to non-scientist audiences, and needed to create visualizations beyond what was anticipated. Non-scientists with whom the scientists were interacting were typically local, state, or national government decision makers involved in developing policies to address environmental problems. The activity of the scientists presenting different kinds of visualizations to non-scientists required the development team to modify both software specifications and development priorities. These new requirements involved more extensive review and revision of the underlying scientific and social problems than expected, and altered some
underlying technical assumptions that had driven initial software design and implementation. Thus, it became apparent that, even if visualization and other innovations were successful in providing scientific insight and solving scientific problems, the challenge of addressing pressing environmental and societal problems – especially those classified as “wicked” – would remain [1].

The late 20th century brought a host of grand challenges to the discipline of ecology, including the concept of coupled human and natural systems (e.g., Liu et al. [2]), which present as complex, dynamic, and adaptive systems. Problems in these complex systems are characterized as difficult or impossible to resolve because of the fluid and often contradictory requirements for any effective or acceptable solution [3]. These types of problems have come to be known as wicked problems. For example, the complex interdependencies of the issues usually result in the creation of new problems even as we think we’re making progress with the original problem [4]. Climate change has been identified by some observers as a “super wicked problem” [5] because of its further exacerbating features including the fact that time is not costless, so the longer it takes to address the problem, the harder it will be to do so. It is in this atmosphere of increasingly complex problems that ecology and other environmental sciences are trying to find solutions to their grand challenges.

The results presented here describe case study findings related to the question: How do wicked problems change visualization software design considerations in a problem-driven design study? The primary contribution of this paper is the contention that understanding the nature of so-called wicked problems and resulting changes in scientific approaches is critical to developers and procurers of software products that convey scientific results to decision makers and the public. While the Association for Computing Machinery (ACM) engineering code of ethics instructs that “…software engineers…[should] consider broadly who is affected by their work,” stakeholders are typically thought of as the developers and users, as well as those who have to support, deploy, or pay for the software; secondary users are at best thought of as those whose information is ingested into software systems [6]. Another way to think about these “secondary users” is as decision makers who are responsible for making choices and policies based on available information. While engineers have traditionally been taught to eschew policy matters [7,8], recent social science research suggests that technology inevitably has policy implications [9] despite best efforts of scientists and engineers to divorce themselves from policy decisions [10]. Results presented here suggest that developers will produce more effective software if they explicitly recognize the role their artifacts play in decision making, and hence in policy, which is relevant to the digital government community as they develop or select software to study complex scientific problems in which the public holds considerable interest.

2. Problem-driven design approach

Problem-driven design approaches are well-tested [11] and highlight the importance of collaborating with partners who are typical of future users of the software [12]. A well-known approach for problem-driven design of visualization software is the Nested Blocks and Guidelines Model [13]. The nested design model defines four levels of design considerations; analysis in the highest level (problem characterization) cascades to affect design criteria at lower levels, ideally aligning with the problem at hand (Fig. 1).

For example, to solve a problem within a certain domain, scientists might rely on datasets that they analyze using statistics or with a simulation; these datasets constitute the data or task abstraction block, where researchers choose which phenomena to measure [14]. Once scientists collect datasets and perform necessary manipulations, transformations, or simulations, a technique for visualizing the data or
Fig. 1. The Nested Blocks and Guidelines Model is used for problem-driven design and development of visualization software [29].

model is chosen. The domain problem characterization thus affects all design decisions about the resulting visualization. A refinement to the Nested Blocks and Guidelines Model can be used to characterize the outermost level of the nested model – the domain problem – and to design for wicked problems, which are by definition dynamic and difficult to problematize [15].

3. Designing for wicked problems and post-normal science

For the most part, traditional “normal science” [15] is unprepared to answer questions posed by policy makers responsible for managing the complex systems that generate the kinds of problems known as “wicked” [1]. New approaches, what have come to be known as post-normal science [3], are emerging as ways to generate the information needed to make intentional and collective choices to resolve wicked problems. There are three key, interconnected components of the post-normal science model that make it different from other approaches and more appropriate for dealing with wicked problems: (1) Uncertainty is considered more than a technical or methodological issue; uncertainty is accepted as the state of affairs within which decisions must be made. (2) Different approaches are recognized and leveraged rather than assuming a scientific or policy consensus can be found. (3) The group of individuals considered capable of assessing the quality of the results is extended beyond the normal disciplinary peer community to a wider range of experts and knowledge; this new group is then better able to consider the array of risks, benefits, and implications for multiple stakeholders [1].

This relatively new take on developing science, especially in the face of wicked problems, is in contrast to normal science. In normal science, peer communities are typically limited to those experts who can judge the quality of the science; for the most part, these are disciplinarily-trained peers (e.g., environmental scientists, biologists, physicists, geologists). When uncertainty and decision stakes increase, the post-normal approach suggests that the peer community can and should be extended to non-disciplinary experts, those with experiential, context, or local expertise. Single scientific discipline and strictly scientific knowledge are, by definition, incapable of capturing the full complexity of such problem settings. In post-normal approaches, the peer community is extended to include not only producers of information but potential information users as well. Non-experts can contribute to knowledge production in a variety of ways including helping to frame problems, providing non-scientific information or data, helping interpret data in the local context, and acting as critical reviewers of the output.
Multiple approaches have been created to bring non-experts into both the production and evaluation of knowledge including consensus conferences, which have been used to bring together competing perspectives and values around topics like bioremediation of hazardous wastes [17]; citizen juries, which have been organized to assess the quality of biomedical research [18]; or the introduction of uncertainty guidance to the Netherlands Environmental Assessment Agency that includes ways to consider both quantitative and qualitative metrics of uncertainty in risk assessments [19]. Another approach is the creation or use of a knowledge-to-action-network (KTAN) [20] that brings together a dynamically evolving group of participants who work together to pose and answer questions collaboratively and iteratively, with the goal of creating usable information (or knowledge). One highly visible and credible example that integrates the idea of an extended peer community is the Intergovernmental Panel on Climate Change (IPPC), which can best be described as an interdisciplinary assessment of scientific research to integrate available knowledge for use by policy makers [21]. This extended peer group is not only an interdisciplinary group of scientists studying the problem, but an interdisciplinary group of scientists and policy experts working together to interpret and understand the consequences of the data.

4. Case study methods and analysis

A qualitative case study [22] of environmental scientists involved in a problem-driven software design project was used to inform the development priorities for the software team developing data visualization tools. These tools would eventually support environmental scientists’ exploration of data and communication of complex datasets. The case study query included interviewing key informants (n = 7) to test the use of visualization as a way to analyze and communicate research output. Each key informant was formally interviewed at least once with a semi-structured interview (Appendix A). Interview questions were based on a number of propositions, including the hypothesis that visualization would help scientists understand their data and come to insight more easily. Developing questions and analyzing responses based on initial propositions are common techniques for framing a case study [22].

In addition to the transcripts from interviews with key informants, data were collected in the form of (1) field notes from emails and informal conversations with members of the group; (2) weekly project meetings; and (3) audio recordings of annual all-hands meetings (n = 21 participants, including the seven key informants). General e-mails and the project website, though not used in the analysis, became an archive of the timeline of events and project highlights. Project meetings varied topically to include both high-level ecology problems and technical detail. All data from interviews, conversations, meetings, and memos were transcribed for an analysis that used themes built from the project’s propositions.

NVivo, a software tool, was used to analyze project data. Primary codes for analyzing the transcripts were taken from keywords in the interview questions, and statements from transcripts were organized into one or more of the following categories for further analysis: (1) exploration, (2) communication, (3) technical challenges, (4) design and development process. Once statements were categorized, frequent themes were defined using key terms and ideas found in the various statements. For example, key informants’ statements about the topic of communication included subtopics such as (1) communication with (a) scientists and (b) non-scientists; (2) the appeal of visualization to various audiences; and (3) design considerations based on audience. Finer grained insights, then, were derived based on statements about these topics and subtopics. Findings presented here include study results involving statements about communication with non-scientists and the design considerations based on perceptions of audience needs.
5. Findings

The environmental scientists involved in the study report presenting their work to non-scientist community leaders on wicked problems related to the impact of climate change. In climate change management questions, the decision stakes rise for policy makers and citizens due to controversy and uncertainty as global models are downcaled to regional areas, and as forecasts are stretched into an unknowable future. This combination of high systems uncertainty and high decision stakes suggests that a new problem solving strategy grounded in post-normal science may be appropriate for developing the information needed to move forward on local and regional decisions. What makes post-normal science unique is that scientists, such as the environmental scientists involved in the visualization software project presented here, work together with non-scientists to co-develop sophisticated models to explore together the impact of climate change in the local context. According to the environmental scientists queried in this study, scientists and community members co-developed the inputs and assumptions of the ecological model used for revealing management options. The visualization software developers were challenged to design software to assist environmental scientists in what they viewed as priority developments. The following findings provide deeper insight into problem-driven design objectives when considering wicked problems, such as climate change, especially the surprising finding that environmental scientists prioritized visualization design that appealed to non-scientists for the purpose of discussing and interpreting model output. The findings here emerged as a result of studying the problem-driven design objectives related to visualization needs of non-scientists audiences.

5.1. Design objective 1: Visualization of topography for realistic storytelling

The original and primary design objective for the software development team involved creating visualizations of physical terrain from flat two dimensions (2D) to incorporate topography for the purpose of scientific exploration. Innovations in data acquisition have enabled scientists to collect data, conduct experiments, and validate models in topographically complex areas. As discovered during case study interviews and meetings with the environmental scientists in the study, topography generally serves different purposes for scientist and non-scientist audiences. The environmental scientists believed that topography appeals to certain audiences because it is an intuitive view of the landscape, grabs attention of viewers, and is photorealistic. This theme emerged again and again over the course of the project in both the interview and meeting data when the environmental scientists distinguished between the different audiences to whom they present visualizations, and described how they would design visualizations with a particular audience in mind. In addition to exploring the science problem of determining the effect of topography on ecological response variables, topography helps non-scientists recognize and relate to familiar landscapes; the third dimension allows secondary users to more easily recognize features on the landscape. Other visualization needs related to portraying topography for non-scientists audiences were to produce animations, (i.e., include a fourth dimension (time)), and to display changes in landscape via fly-throughs that highlight particular areas.

5.2. Design objective 2: Visualization for simplifying the effects of complex interactions between systems

In addition to generating fly-throughs for a stakeholder decision process, one environmental scientist described the need for better understanding how biophysical systems, management actions, and socio-economic influences interact to affect sustainability in fire-prone landscapes under climate change.
The audience for whom visualization would be designed included state and federal agency staff, non-governmental organization members, local and county officials, university extension agents, canal company representatives, and area residents interested in thinking about the future of the region. In addition to the use of visualization, other stakeholder engagement activities were being used including backcasting [23], concept mapping, and a system dynamics model of the hydrological system [24]. The environmental scientist reported on visualization needs for communicating in concert with these other engagement activities. Visualization would be used to simplify complex interactions to show their impact over time.

The above activities deepened all participants’ knowledge of relationships among natural and human systems, and others’ perspectives and values around those systems. The concept map exercise moved the group forward in identifying variables they wanted to consider, data they would need to support those variables, and what probably could not be included in the model due to limitations in data or research methods, and brought about two impacts critical to the software design effort: (1) Non-scientist participants realized that a complex model was critical to understanding and mitigating the problem, and (2) scientist participants tasked with developing that model were provided with assumptions and variables that the community viewed as important and that would drive the model. In other words, the process identified gaps in knowledge for both non-scientists and scientists.

As a result of these reported activities, software developers were tasked to help create visualizations that could help explain and simplify the results generated by the complicated model to answer questions about the impact of climate change in this western US basin. Thus software development of visualization was expanded to serve non-scientist audiences. These new features would challenge developers to find ways to display model assumptions and levels of uncertainty and aggregate variable types to simplify visualizations, as well as increase audience attentiveness with 3D maps, animations, and fly-throughs (Design Objective 1).

5.3. Design objective 3: Visualizing human constructs (boundaries) on the landscape

Researchers from a federal agency sought to demonstrate results from their eco-hydrological model for use in community and regional decision making [25]. Given a set of drivers (e.g., temperature, precipitation) and disturbance (e.g., fire, harvest, fertilization), the model reveals the interaction of stream flow and biogeochemical processes, and carbon and nitrogen dynamics in plants and soils. Running on a daily time step across thousands of pixels, the model generates multiple gigabytes of output for multi-century simulations of large landscapes. The model results are difficult for scientists to tune, interpret, and communicate without visualization.

Unanticipated specifications for the scientific model arose as the environmental scientist presented the results of model visualizations to non-scientist stakeholders. For example, the environmental scientist found it easy when refining his science model to view one image with land use then to imagine in his mind’s eye the land use boundaries on visualizations of nitrate flux, but the non-scientist audience wanted to see the land use boundaries explicitly. This design goal based on the needs of the non-scientist audience was critical enough for the environmental scientist to re-prioritize development goals. While this seems like a simple change in the visualization, for technical reasons it required the software development team to re-think how to render the overlay in the visual output. Achieving this feature involved re-implementation of the underlying graphical rendering technique, a non-trivial undertaking.

Visualizations created by the environmental scientists become artifacts viewed by secondary users, who for the most part are not scientists. This finding aligns with the post-normal science paradigm: the extended peer community often requires the scientist to design visual output to enable others’ understanding.
6. Discussion

In this problem-driven design study, exploring the context of the wicked problem that is climate change turned out to be important, as scientists and non-scientists alike grappled with characterizing the scope of inter-connected environmental problems as they changed over time and space [26,27]. The results presented here do not advocate that scientists and engineers practice normative science, but that they become more aware of the information, technology, or science needs of those who set policy and make decisions. When the decision stakes are high, other factors in addition to information access, such as values and trust, are likely to sway public opinion [28]. In light of recent studies in public policy and philosophy of science that suggest that policy and government decision making are not greatly influenced by scientific research results and, conversely, that non-scientists rarely influence the formulation of science problems [9], it seems critical to extend both the scientists’ and software engineers’ understanding of both primary and secondary users of their science and technology.

Our findings suggest that scientists, software engineers, and systems analysts who co-design and co-develop such artifacts should be cognizant of the type of problems to which those artifacts will be applied. Development teams using a problem-driven approach, then, should understand the characteristics of wicked problems during their design process, which increases the design objectives to include both primary and secondary users of visualization and related software. In traditional scientific visualization projects, computer scientists often suggest and prototype technological innovations to enhance the scientific content of the visualization. However, few computer scientists are trained in which scientific visualizations work for non-scientists in a wicked problem context. Similarly, visualization research studies often focus on tasks, techniques, and algorithms, or lower-level software design considerations in order to boost automation of tasks and the power of machine learning [29]; whereas, findings of the case study presented here show how a problem-driven design process can create new challenges and opportunities for software development teams as they respond to a dynamic problem and decision environment.

According to case study interview and meeting data, stakeholder audiences are characterized as trusting of visualizations that are relatively familiar or intuitive to what they hold in their minds. In other words, for certain audiences, matching what they have in their minds to what they see in the visualization is often proof of truth or fact (e.g., Design Objective 1). While scientists are trained to ward against this kind of confirmation bias, other stakeholders may not be so trained. On the other hand, all viewers might be critical of results if they encounter a visualization contrary to what they expect or intuit. One environmental scientist in the study alluded to this problem when discussing the concept of mental models during a discussion of data exploration practices. He commented on instances when scientists used visualization to validate their data, only to find out that something was wrong with the data. In such instances, the visualization does not match the viewer’s mental model of how the data should look. When a relatively uninformed audience has the same experience of not seeing what is expected, they might also question the data, methods, or visualization, uncovering problems that escaped the trained viewer. More research on best practices for communicating via visualization seems merited based on this analysis.

Design Objective 2 – simplifying complex systems interaction to increase understanding – has broad implications for the importance of developing powerful tools for policy/science collaborations. Findings in other research in the e-government field suggest that software tools are increasingly used within government agencies to enable collaboration between scientists and non-scientists including tools for supporting stakeholders in environmental decision making [30], decision support systems that bridge
science and values [31], and information systems used in a mediating role for tackling climate change adaptation [32]. These studies complement an increasing interest within e-government research on using open data and visualization not only to improve government efficiency but also to make closer connections between citizens and government [33] and to enable stakeholders themselves to make sense of data [34]. Our case study findings suggest that scientists’ use of visualizations to both explain complex results to stakeholders and to engage them in knowledge-to-action-networks could increase trust between government-employed or -funded scientists and decision makers; whether this would pave the way for higher use of e-government services remains, for us, an open question [35]. Visualization developed for such a purpose will need to be designed to present an accurate but simplified data story about complex systems and the related data.

Design Objective 3 highlights the importance and difficulty of human constructs, such as boundaries, which might require scientists to include policy implications as part of their data storytelling using visualization. The visualizations created by environmental scientists in the case study were seen as artifacts to be viewed by secondary users. The related realization that emerged over the course of the research is that visualizations enhance scientific insight not only for stakeholders but also for scientists during collaboration in decision making. Delivering science into policy and managerial processes can challenge scientists; the environmental scientists in the study presented here prioritized visualization tools to equip them to better communicate and explore new findings both in their own work and in broader policy processes with a host of new requirements valued by stakeholders, such as boundary lines.

Key to designing both the tool and the visualization is the realization that scientists are most likely to use visualization when telling a story about the results of their scientific research. More generally, research on group interaction with visualization, rather than single-user interaction, could provide insight into the contribution of visualization to the process of communicating results, or of telling a data story. And, finally research into how to create visualizations for non-scientists – which visualizations work and why – is needed. Training for both computer scientists and scientists on designing such visualizations for diverse audiences seems merited based on the findings presented here.

7. Conclusion

Throughout the software design process, a problem-driven approach emphasized close collaboration with primary users. While the ecological problems had been the primary focus for developing visualizations, case study research showed that in certain settings where non-scientists were involved, generating suitable visualization was considerably more complex than originally anticipated. Computer scientists and environmental scientists involved in design and implementation of both visualization tools and visualizations for wicked problems must take particular care to fully characterize the domain problems early in the design process, and throughout the subsequent implementation, deployment, and maintenance of software tools.

The primary lesson of the case study presented here is that the problem domain affecting the design of the visualization is likely much broader than originally conceived. Software developers working with clients, such as the environmental scientists in this study, should be prepared to recognize wicked problems and be cognizant of the range of audiences interacting with tool output, such as visualizations produced via a software tool. Finally, while these considerations will help improve design decisions made at the beginning of the project, it is unlikely that all decisions will be correct as the project matures, and it might become necessary to revisit software design decisions as they emerge. Developers and
users need to be prepared for the time and cost of revising assumptions that drive initial technology decisions. These findings have great benefit for research and development teams involved in similar software design and development projects where scientific data is to be visualized and used in decision-making processes and public policy settings.

Future research for similar projects might include testing the output of the visualization tools with various audiences in different settings. Additionally, researchers might track and measure the role of visualization in a problem-solving process, especially the extent to which the audience can use the visualization to provide insight into a problem, and where visualization might fall short of that goal.

Acknowledgements

VISTAS is funded by NSF BIO/DBI 1062572. The authors would also like to acknowledge the entire VISTAS team, including senior scientists Mike Bailey (Oregon State University – OSU), and Jenny Orr (Willamette University), and developers Nik Stephenson-Molnar (Conservation Biology Institute – CBI), Taylor Mutuch, Viriya Ratananasapunth and Chris Schultz, and our collaborators Bob McKane and Allen Brookes (U.S. Environmental Protection Agency), John Bolte (OSU), Christoph Thomas (Bayreuth University), and Dominique Bachelet (CBI). We also thank anonymous reviewers, and of course Jing Zhang (Clark University) and Yushim Kim (Arizona State University), for helpful suggestions.

References

Appendix A

Semi-Structured Interview Protocol

1. Tell us a little bit about the work that you do, especially the work that involves large data sets.
2. Are you familiar with visual analytics or visualization processes? If no, go to Question #3. If yes:
   – Are there aspects of your work that could benefit from visual analytics or visualization?
   – Are you using visualizations now? If no, go to Question 3.
   – Can you demonstrate or show us some of the visualizations you use now (5 maximum)?
   – How do you acquire the visualizations you currently use? Do you create them yourself or is the work contracted out to a third party (programmer, contractor, student)? Is the visualization software directly attached to your model? How much of your own time (or money) do you invest in acquiring these visualizations?
   – What purpose does each visualization serve?
   – How well do these visualizations achieve those purposes?
   – What do these visualizations do well?
   – What are the shortcomings (if any) of these visualizations?
   – Were you able to communicate what you wanted to accomplish?

[References]

3. How helpful do you think visualizations could be/are for you in your research efforts?
   - In framing problems?
   - In creating new hypotheses?
   - In exploring data?
   - In understanding data?
   - In communicating research results with other scientists in your discipline?
   - In communicating research results with others outside your discipline?

4. What are you expecting to learn from participating in this project?

5. What difficulties are you expecting while participating in the project?
   - Communicating with/understanding the computer scientists working on the project?
   - Communicating with others outside discipline?
   - Understanding how the visualizations work? Increasing understanding of the underlying ecological concepts?

6. What visualization(s) and/or visual analytics would you especially like to explore during this project?

7. Anything else you’d like to share regarding visualization, this project, your participation?