Using Visualization to Derive Insights from Research Funding Portfolios

Andreea Molnar
Portsmouth University

Ann F. McKenna and Qing Liu
Arizona State University

Mihaela Vorvoreanu and Krishna Madhavan
Purdue University

Over the past few years there has been increased interest in open data access. However, data retrieval represents a major bottleneck for many organizations, including funding agencies such as the US National Science Foundation (NSF). This is due to the multimodal, high-volume nature of funding data that is often disconnected and diffuse. In effect, deriving meaningful insights from an existing funding portfolio is challenging. Issues arise not only in data representation, but also in the way that data is accessed and communicated to the end user. Providing multiple ways to visualize data, enabling users to "see" relationships, gaps, or other connections, could lead to new insights that facilitate a better overview of existing funded research and more informed decision making for future funding.

As with any federal funding agency, there is complexity of data within NSF's portfolio, and it increases continuously as new proposals get submitted, reviewed, and funded or declined. NSF awards approximately 11,000 grants a year with an average performance duration of three years. The NSF staff includes 1,400 career employees, 200 scientists from research institutions on temporary duty (rotators), and 450 contract workers and staff (data obtained from www.nsf.gov). Given the relatively high percentage of turnover among temporary NSF staff, there are always new individuals that need to be trained on funding processes. It is crucial for them to familiarize themselves quickly with the funding portfolio. For example, science assistants are typically hired for a period of two years and are expected to help program officers with queries on the data using various criteria. Because information increases continuously, and the staff need to make decisions in a short period of time, there is a need to make the existing portfolio easy to access and understand.

In this environment, however, the gap between data visualization and actual insights can be difficult to bridge. To address this problem, we developed Deep Insights Anytime, Anywhere (DIA2), a data-mining and Web-based platform that uses metadata about funding information from the NSF to produce interactive data visualizations.[1] DIA2 provides an overview in an easy-to-understand format, allowing users to search, view, and analyze the NSF funding portfolio.

This article presents an assessment of DIA2's usability. We attempted to evaluate its perceived utility among the target user group. Specifically, we inquired whether and how the system leads to insights that can facilitate further decision making, such as estimates about the impact of specific research projects. Our results show that DIA2 has good usability. Furthermore, participants identified several indicators of impact as a result of the visualizations that can be realized through DIA2.

DIA2

During the design and implementation of DIA2, we followed an iterative, user-centered design process.[2] To determine DIA2 requirements, we

 0272-1716/15/$31.00

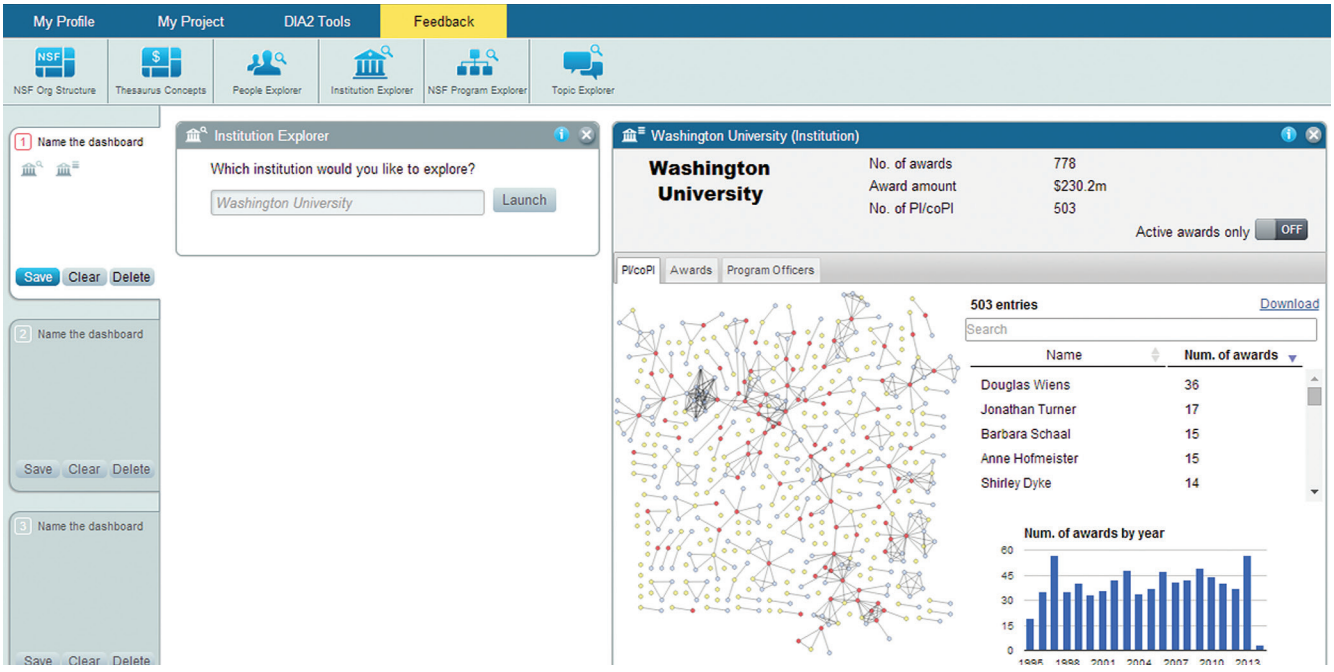


Figure 1. DIA2 search functionality: The search results show awards at Washington University. The institution widget displays the search results across multiple tabs.

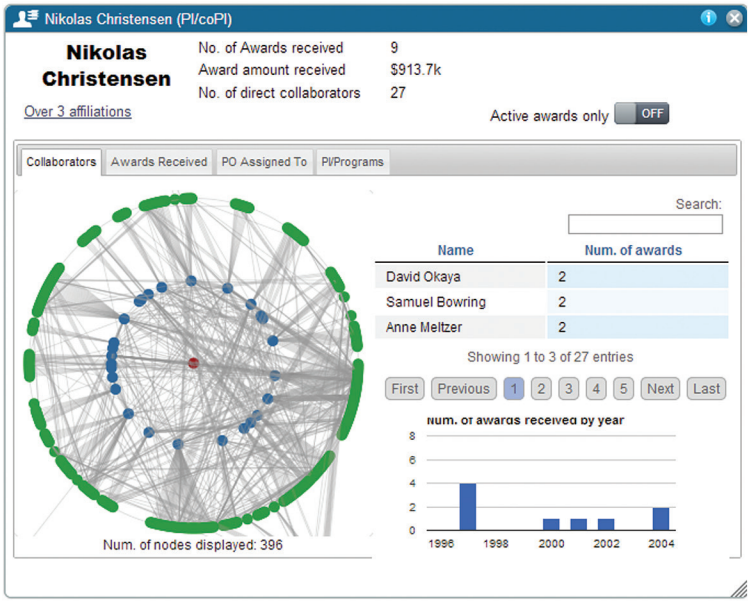
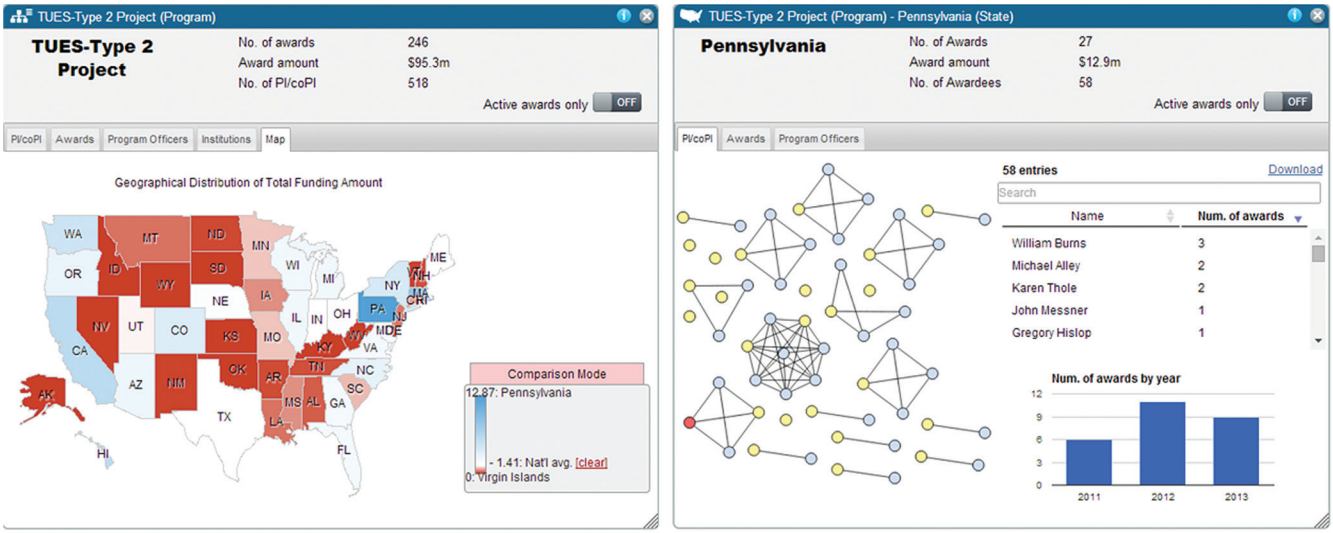


Figure 2. Investigator-centric (ego-centric) network: This type of network is used to show direct and second-level collaborations. Colors convey similarity between nodes (maroon represents the researcher, blue are the other researchers directly connected to the researcher, and green shows second-level collaborators).

conducted interviews with NSF staff and distributed user diaries to participants to collect their daily work patterns and thoughts as they used the DIA2 system.[2] This led to a model of user profiles, workflows, and requirements that informed features of the DIA2 design and implementation.

The DIA2 search functionality and visualization services are presented in widgets using tabs that each present a certain characteristic of the search results. For example, as Figure 1 shows, when a search is performed for NSF funded projects at Washington University, the information is presented across three tabs. The first PI/co-PI (principal investigator/co-principal investigator) tab shows a network representing collaborations among the funded PI/co-PIs, with an embedded table that includes the PI names and total number of awards received. The second tab presents the awards made, and the third tab shows the program officers who are managing these awards. As the icons at the top toolbar in Figure 1 indicate, DIA2 lets users search for people, institutions, programs, and topics. These search capabilities were determined based on activities, frustrations, and data needs expressed in user interviews, focus groups, and diaries.[2]

In addition to a hierarchical data visualization service and a geographical data representation service, DIA2 offers a collaboration data visualization service that focuses on two types of collaboration graphs. The first one is used to visualize the researchers' collaboration networks through the use of a spring-loaded social network layout (a spring-based layout where a spring represents a collaboration and a vertex a researcher), as Figure 1 shows. This graph type allows aggregation of collaborators at the institution, program, or specific topic levels. The other type of collaboration graph, shown in Figure 2, is an investigator-centric (ego-centric) collaboration network showing direct collaborators and second-level collaborators (or collaborators of direct collaborators). Colors are used to convey



(a) (b)

Figure 3. DIA2 geographical data consolidation: (a) Map-based geographical visualization. (b) Selecting a certain region displays the details for the organizations that received funding.

similarity between nodes: maroon represents the researcher, blue are the other researchers directly connected to the researcher, and green shows second-level collaborators.

In both the Figure 1 and 2 graphs, people are represented as nodes, and edges indicate a link between two researchers. The investigator-centric network primarily displays all the collaborators a person has worked with on NSF proposals and second-level collaborators. This allows users who are looking for collaborators to determine if there are any people who overlap between their connections and the potential collaborator connections that could facilitate an introduction. Both collaboration graphs combined provide end users with different levels of granularity for identifying collaborators.

DIA2 also provides geographical data consolidation that illustrates how NSF funding is distributed across geographical areas (see Figure 3). DIA2 uses maps to provide a visualization interface that supports information access and exploratory activities. The interface allows geographical data to be visualized in *comparison mode*, in which the data is compared with the national average, or *standard mode*, in which the intensity of the color represents the amount of funding awarded to that region. This allows NSF staff to determine how funding is distributed over different states.

Various elements of the DIA2 widgets let users explore the funding portfolio interactively. For example, if the user clicks a node (a researcher or PI/co-PI) from the graph showing the project collaborators, that will open a new widget that provides more details about the selected person. Similarly, clicking on a thesaurus concept triggers another widget to open that provides an overview of the funding outcomes related to the given concept.

A previous study evaluated the usability of individual widgets and the data visualizations within them.[3] Here, we focus on the system's overall perceived utility. Although usability is one necessary component of utility, it is by no means sufficient. We aim to validate that users are able to indeed derive meaningful insights from DIA2.

Evaluation

In an effort to evaluate the system's overall perceived utility, we addressed the following research questions: How would NSF staff use DIA2? How would NSF staff use DIA2 to estimate the impact of specific research projects?

Methodology

This study took place either at NSF headquarters or virtually over the Internet. During the study, participants were asked to use DIA2 to perform different tasks, think aloud while performing the tasks, and interpret the data visualizations. The research protocol included a list of tasks to be performed, but the participants were free to explore different aspects of the applications if they wished.

The study took approximately one hour to complete and consisted of three parts. In the first part, participants were asked to perform four tasks that involved finding specific information using the DIA2 application. Participants were then asked to interpret the results. This part helped users familiarize themselves with the system and provided us with data about how participants make sense of the various data visualizations.

In the second part, users were asked to provide examples of tasks they need to perform in their daily activities. This information was requested to help us understand whether DIA2 could be used to perform the tasks the user identified. If researchers determined that DIA2 could allow the successful completion of the user-identified tasks, the user was asked to proceed with them. This was followed by several questions regarding users' perceived interest in using DIA2 in the future and in what context it would be useful. Questions regarding further improvements of the application were also asked at this stage. In both the first and second stage of the study, we asked questions relating to indicators of impact as presented in DIA2.

The third part of the study consisted of users completing a questionnaire. The aim was to quantify users' perceptions regarding the effectiveness of DIA2's usability and design and to collect basic demographic data.

Separate from the suggested areas for future development, we obtained substantial feedback on aspects of the interface design and functionality.

Recordings of participant interactions were made either using a video camera (for face-to-face data collection at NSF) or via screen and audio capture using WebEx for remote data collection. During the study, notes were taken by the facilitator for later analysis. The study recordings were transcribed later, and the transcripts and notes served as data sources. During data analysis, we identified codes, or patterns, which we then grouped into categories.

Participants

We used a snowball (referral-based) sampling technique[4] to select participants, but we also aimed to have a representation of NSF staff that would benefit from using the tool. Twelve participants took part in the study, aged between 25 to over 65 years old, with more females (64%) than males. We were mindful to include two crucial and interacting roles: the frontline analysis (the domain expert doing the actual data analysis) and the gatekeeper (the person who authorizes the data usage).[5] In our case, the frontline analyses were typically done by science assistants at the request of program officers (the gatekeepers).

User Interaction with the System: Perceived Utility

User interaction data indicated positive aspects of DIA2 as well as areas for further refinement. For example, participants expressed the need for additional functionality such as the ability to

- export data ("Frequently, program officers like to tweak the data we give them, so we have to give them a form that they can do that in, and usually it's Excel.");
- zoom in and out of graphs, especially for the graphs where many nodes are displayed ("Would be nice to be able to have a little bit of the zoom feature, so I don't have to do a massive zoom, but I can do one stage of zooming in around a particular cluster.");
- rank the programs by dollar amounts;
- visualize the program and collaboration evolution over time to see a program's predecessor;
- obtain email addresses for the funded PI(s)/co-PI(s) ("A lot of the questions that we get as science assistants are, what are the email addresses for these PIs? A lot of times we have to send all active awardees for this particular program an update if there's a PI meeting coming, or in the case of, with this new IUSE [www.nsf.gov/funding/pgm_summ.jsp?pims_id=504976] program, don't worry, even though you have a TUES [www.nsf.gov/funding/pgm_summ.jsp?pims_id=5741], you're still going to be funded. That type of thing."); and
- access other information about a project that may be found in the public domain (such as reports, presentations, and publications).

Separate from the suggested areas for future development, we obtained substantial feedback on aspects of the interface design and functionality. For example, users made several positive comments about how the information is represented. There was general agreement that the user interface was intuitive and easy to navigate and that the visualizations were effective in representing information.

One of the participants particularly praised the ability to open two widgets at the same time and compare the information displayed: "I like that you can compare [two institutions]. That's neat!" We also found positive comments on the ability to sort individuals based on the funded proposals—"I like that if I go like this, I can automatically sort it"—and the ability to filter information displayed on the widget based on the keywords entered by the user in the search box. Others also expressed that DIA2's functionality

was better than the existing tools they currently use at NSF: "this [Program Explorer widget] is really cool for me to see the connections [PI/co-PI collaboration network] and it's unlike anything I could do before."

Most of the users reported their intention to use DIA2 in the future. For those who had access to DIA2 prior to our study, they provided examples of how they had already used DIA2. For example, one of the participants reported that s/he "would totally use" DIA2 in his/her daily activities, and another reported having already used some of the features provided and "because this is such a neat [collaborators graph] we wanted to use it for our PI meeting … I think the community was really impressed by that because it shows a lot of the connections." Another participant had used DIA2 intensively in trying to better understand the portfolio: "a majority of what I've used it for … is looking at programmatic activities to … see what the portfolio may represent."

The topic explorer feature was mentioned by program officers as useful for analyzing funding at different educational levels. "Any time a topic comes out of the front office I'll immediately punch it in here to be able to analyze what may take place.… Does K through 12 look very different than undergraduate versus graduate versus informal?"

Overall, responses and feedback indicated that the system provides useful visualizations and that the tool has an effective interface. Besides DIA2's overall utility, we were also interested in understanding how participants could use the tool to derive actionable insights.

Metrics of Impact

Funding agencies often do not have the tools to advance understanding of research impact or provide metrics for measuring success.[6] Thus, in addition to obtaining user data to inform the design and functionality of DIA2, we were also interested in whether and how users interpret the visualizations in the context of defining impact. The coded data was grouped into three main categories: PI/co-PI, institutional, and specific NSF program impact.

Responses indicate that the number of awards received could be relevant in highlighting the impact of a researcher: "If they were given several awards I would think that they're making a big impact. I wouldn't think they would be getting awards if they weren't." Another metric offered was the amount of money received because it indicates that NSF is "voting with money the value of that research moving forward." The participants also commented that the number of collaborators and their collaborator networks are another indicator of impact because "the size of the collaborative network is going to end up as a measure of NSF investment … over time."

Participants reported that the consistency with which an institution received funding could be another indicator of impact. One participant mentioned clustering of the nodes in the network would help users explore whether there is impact at the institutional level, as a way to explore whether there is "some cross-fertilization taking place." The participant suggested further that, by comparing two institutions in terms of impact, one institution may have been more successful than the other because the institution was "able to explore research opportunities in more interdisciplinary and collaborative ways."

Overall, responses and feedback indicated that the system provides useful visualizations and that the tool has an effective interface.

Participants also stated that one way to look at the impact of NSF programs could be the number of awardees. Another way to measure success is to look at the program distribution across states—that is, if the funding is distributed across several US states.

Researcher, institution, and program were three main areas mentioned by the participants as potential criteria for determining impact. The participants also suggested ways in which DIA2 could integrate other functionality for assessing impact at each level. For the individual level (PI/co-PI or researcher), users suggested that DIA2 could integrate the researcher's publications and citations. At the institutional level, users suggested that a visualization of the leadership group and quality of research could be useful. At the program level, they suggested integrating the number of awards that continued to be funded, the number of students, and the number of faculty that have been impacted by the research funded in a certain program.

Some participants pointed out that this is neither an exhaustive nor necessarily representative way of measuring impact: "I don't feel comfortable with this—from either of these representations of being able to … of us being able to identify what impact is." Others highlighted that impact

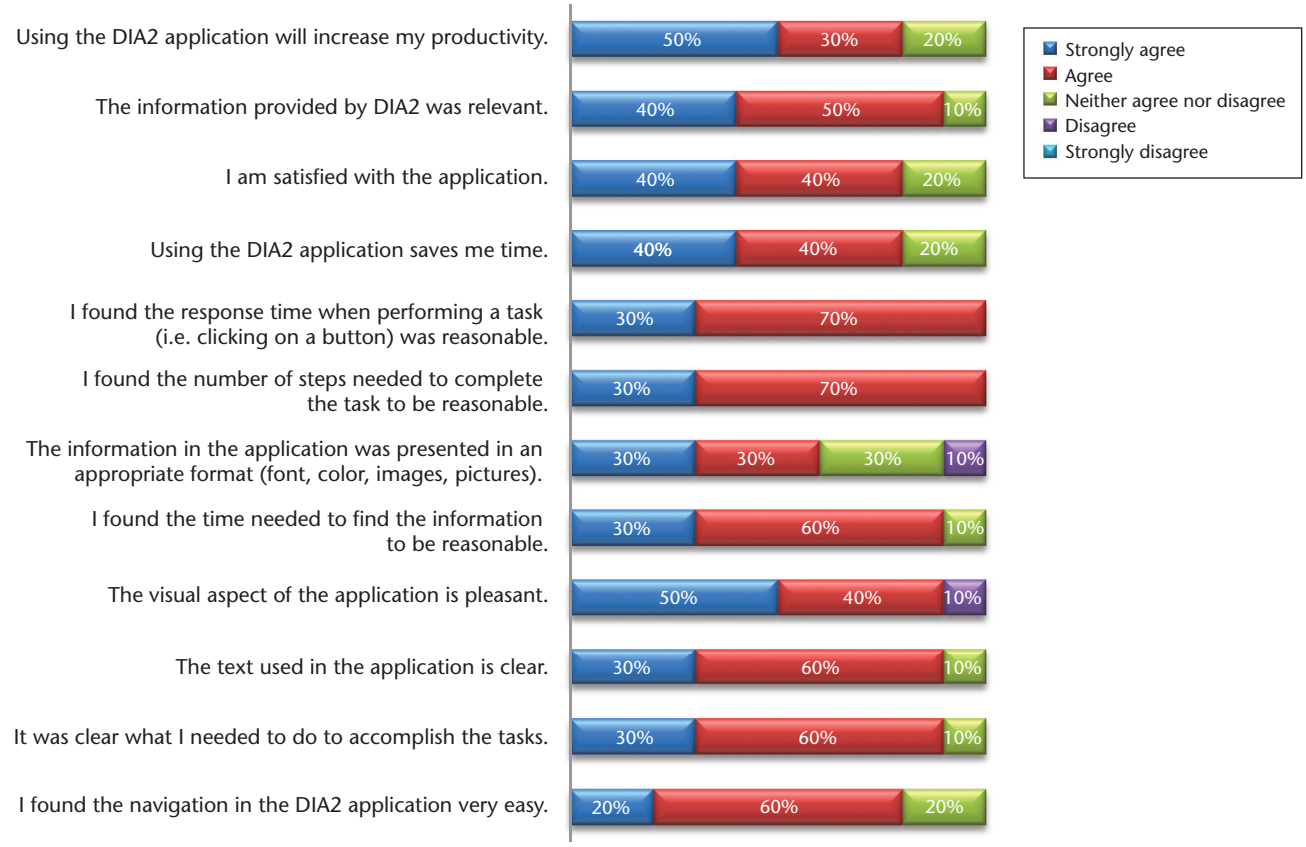


Figure 4. DIA2 system perceived utility questionnaire results. Participants responses to questions targeted at specific aspects of DIA2.

depends on the context used, and it cannot be measured over a short period of time: "I think impact—how I define impact is going to be different for different directorates and for different divisions. Some of it is longitudinal. We're not going to know for years. Some of the stuff that's been funded ..., we won't know for 20 years whether it really had impact, but certainly beginning to look at some of the collaboration, etcetera, will help us get a sense of what's important." Some of the participants highlighted the broader impact on society: "Is it going to change the world?" Moreover, in educational settings the sentiment is that it is important that the funding has made an impact on student experience/learning or teaching practices: "ultimately what we're concerned about is really the number of faculty and students who are reached, and that's a hard thing."

Results from this study provide a starting point for how DIA2 can be effective as a tool for defining and measuring impact and also show that impact of research is complex and difficult to operationalize. Additional work is planned to collect data from a more comprehensive group of users, particularly participants from the research and teaching communities (not just within NSF). We hope that this will broaden the perspective on DIA2 and help to further characterize impact.

Questionnaire Results

In addition to performing user studies and interviews, we also collected quantitative data based on the System Usability Scale (SUS),[7] which is a short questionnaire consisting of 10 items that assess perceived usability. SUS is reliable and valid,[8] is technology agnostic, and has been extensively used to assess the usability of applications in different contexts. SUS makes it possible to quantify usability and thereby compare systems. Based on data from the current study, the DIA2 SUS score was 78.33, which is higher than the average usability score of 68.[9]

SUS is not intended to be a diagnostics tool.[7] We, therefore, added questions that could shed light on perceptions of specific aspects of DIA2. Figure 4 shows the participants' responses. Overall results indicate that the DIA2 interface is visually appealing, the information it provides is relevant, and the navigation is easy. Users also indicate that DIA2 increases productivity, thereby indicating its utility as an application.

During the user studies, participants did note some confusion about colors and the placement of links. However, for two items in particular—"I found the number of steps needed to complete the task to be reasonable" and "I found the response time when performing a task (i.e., clicking on a

button) was reasonable"—all participants chose either agree (70 percent in each case) or strongly agree (30 percent in each case). These results are encouraging, with most of the participants providing positive feedback on DIA2's functionality, usability, and utility.

Our results indicate that DIA2 could be seen as a complementary source of information that can be used by researchers, universities, and funding agencies to visualize and determine measures of impact that could help inform strategic choices. The current study also revealed areas for system improvement such as providing the ability to export data as well as zooming features on collaboration/network maps. This is a known limitation of the node-link diagrams, which become unreadable when they reach a few thousand nodes.[10]

DIA2 has been refined based on the feedback obtained as a result of the study. The user community considered in this research consisted entirely of NSF staff. Although we obtained positive feedback from NSF employees, it was necessary to get input from PIs, university faculty, college teachers, and other community members to understand the impact and usability of DIA2. User studies with the community have been performed and the results are currently under analysis. Based on these results, DIA2 will be enhanced to provide functionality necessary for the broader higher education community.

Acknowledgments

This work was supported by the US National Science Foundation awards TUES-1123108, TUES-1122609, and TUES-1123340. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the NSF.

References

1. K. Madhavan et al., "Portfolio Mining," *Computer*, vol. 45, no. 10, 2012, pp. 95–99.
2. Q. Liu et al., "Designing Discovery Experience for Big Data Interaction: A case of Web-based Knowledge Mining and Interactive Visualization Platform," *Design, User Experience, and Usability. Web, Mobile, and Product Design*, LNCS 8015, Springer, 2013, pp. 543–552.
3. K. Madhavan et al., "DIA2: Web-Based Cyberinfrastructure for Visual Analysis of Funding Portfolios," *IEEE Trans. Visualization and Computer Graphics*, vol. 20, no. 12, 2014, pp. 1823–1832.
4. L.A. Goodman, "Snowball Sampling," *The Annals of Mathematical Statistics*, vol. 32, no. 1, 1961, pp. 148–170.
5. M. Sedlmair, M. Meyer, and T. Munzner, "Design Study Methodology: Reflections from the Trenches and the Stacks," *IEEE Trans. Visualization and Computer Graphics*, vol. 18, no. 12, 2012, pp. 2431–2440.
6. J. Lane and S. Bertuzzi, "Measuring the Results of Science Investments," *Science*, vol. 3331, no. 6018, 2011, pp. 678–680.
7. J. Brooke, "SUS: A Retrospective," *J. Usability Studies*, vol. 8, no. 2, 2013, pp. 29–40.
8. A. Bangor, P. Kortum, and J.A. Miller, "The System Usability Scale (SUS): An Empirical Evaluation," *Int'l J. Human-Computer Interaction*, vol. 24, no. 6, 2008, pp. 574–594.
9. J.A. Sauro, *Practical Guide to the System Usability Scale: Background, Benchmarks, & Best Practices*, Measuring Usability, 2011.
10. C. Dunne and B. Shneiderman. "Motif Simplification: Improving Network Visualization Readability with Fan, Connector, and Clique Glyphs," *Proc. ACM SIGCHI Conf. Human Factors in Computing Systems*, 2013, pp. 3247–3256.

Andreea Molnar *is a lecturer in the School of Creative Technologies at Portsmouth University, UK. Contact her at andreea.molnar@port.ac.uk.*

Ann F. McKenna *is a professor and director of the Polytechnic School in the Ira A. Fulton Schools of Engineering at Arizona State University. Contact her at Ann.McKenna@asu.edu.*

Qing Liu *is a graduate research associate in the Polytechnic School in the Ira A. Fulton Schools of Engineering at Arizona State University. Contact her at Qing.Liu.7@asu.edu.*

Mihaela Vorvoreanu *is an assistant professor and the graduate program coordinator of the Computer Graphics Technology program at Purdue University. Contact her at mihaela@purdue.edu.*

Krishna Madhavan *is an assistant professor in the School of Engineering Education at Purdue University. Contact him at cm@purdue.edu.*

Contact department editor Mike Potel at potel@wildcrest.com.