

Needs Assessment for the Design of Information Synthesis Visual Analytics Tools

Anthony C. Robinson
GeoVISTA Center, The Pennsylvania State University
{arobinson@psu.edu}

Abstract

Information synthesis is a key portion of the analysis process with visual analytics tools. This stage of work requires users to collect, organize, and add meaning to individual analytical results. This paper reports the results of a needs assessment study with technical and bio/chemical security analysts intended to characterize the ways in which users currently synthesize information, and to elicit ideas for future tools to support information synthesis. Our work used structured interviews to obtain knowledge from analysts. Responses indicate that synthesis is currently supported through the use of office productivity software, and current tools do not provide adequate support for the task of information synthesis.

Keywords--- Synthesis, Needs Assessment

1. Introduction

The process of using visualization has been described by some as beginning with exploration, continuing to analysis, transitioning to synthesis, and concluding with presentation [1, 2]. This process involves a transition from the private realm of an analyst working by themselves, to the public realm where analysts work with others and ultimately share their results (Figure 1). To date a substantial amount of research has focused on exploration, analysis, and presentation tasks facilitated by visualization, but little has been done to characterize and design tools for supporting synthesis.

This paper presents the results of a needs-assessment study intended to characterize the current state of the art of synthesis, and to elicit ideas for future synthesis support visual analytics tools. In this paper, synthesis is defined as the tasks associated with collecting, organizing, and adding meaning to analytical results. In the context of a real-world analysis, an analyst might be faced with the synthesis task of bringing together information chunks that were generated by a variety of visual analytics tools with the goal of developing an overall situational picture (Figure 2). This type of task would happen before a formal presentation of results.

We begin with the motivational context for this work and a discussion of recent development efforts in

visualization that are intended to support the process of information synthesis with visual analytics tools. Next we present results from interviews with analysts to characterize how synthesis is currently conducted and to envision new synthesis support tools. We conclude with thoughts on what these results mean for visual analytics researchers and tool development.

2. Motivation

Motivation for this work comes in part from the need for integrative visualization applications that move beyond support for exploration and analysis and connect those actions to knowledge construction and representation. A criticism of current visualization tools is that they are too data-centric, providing few functions intended to help users develop concepts and higher-level understanding from the results of visual exploration [3]. Therefore it is important to design and implement new visualization tools that bridge this gap.

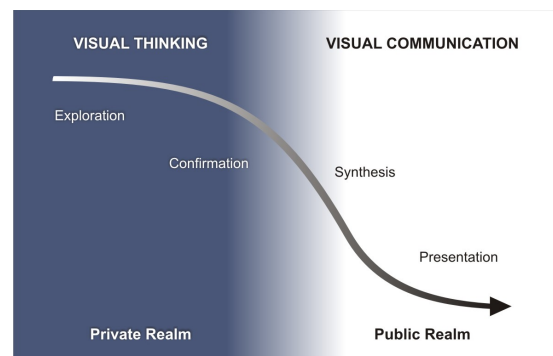


Figure 1 Research process using visualization (after DiBiase, 1990)

Gahegan [3] suggests that users are likely to move back and forth between exploration, analysis, synthesis, evaluation, and presentation tasks in a non-linear fashion. Because it is difficult to support all of these tasks in a single toolkit, visualization tools should instead be engineered to easily coordinate with one another through a common framework. This would allow users to discover patterns in one tool, and pass this information into an information synthesis tool that helps them organize and make sense of these discoveries without

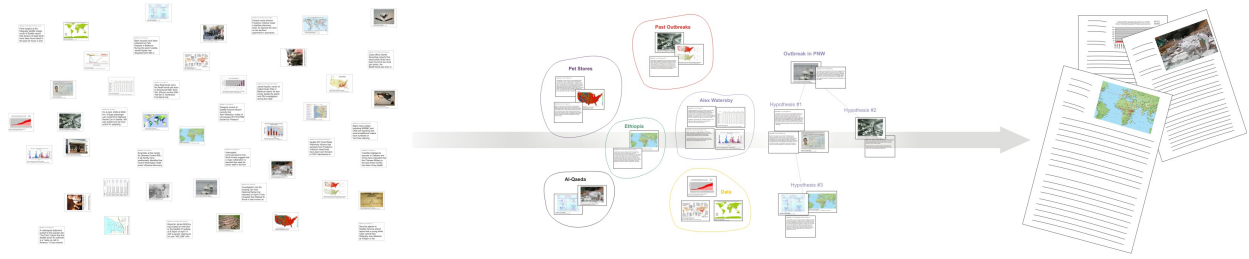


Figure 2 The synthesis process, a transition from individual results to collections, and finally to reports

losing provenance information.

Much attention in recent visualization research focuses on visual analysis of data from diverse sources like text, video, imagery, numerical tables, and spatial information [4]. In that scenario it will be even more important to preserve provenance information in a way that lets analysts browse representations of their knowledge and easily recall prior work. Such systems would allow users to preserve important metadata about their results, and this metadata could include the information necessary to recreate the exact scenario in which they were generated.

3. Background

Visualization researchers have begun to tackle the problem of supporting synthesis with new visually-driven environments for collecting and adding meaning to analytical results. Current synthesis support tools include Analyst's Notebook [5], nSpace [6], GeoTime [7], Scalable Reasoning System [8], EWall [9], and Jigsaw [10]. A common theme driving the development of such systems is that providing users with interactive, visual interfaces for constructing knowledge from analysis artifacts will help them develop compelling stories about their findings that can be presented to decision makers [11].

Analyst's Notebook [5], a tool developed by i2 Inc. is commonly used in intelligence analysis and law enforcement settings to organize and add meaning to collections of information [6, 12]. Analyst's Notebook allows users to assemble multimedia and construct layouts that describe complex scenarios. Its interface provides a large, blank canvas on which users can arrange individual pieces and assign text, photos, or graphics to represent information and the links that may exist between multiple fragments.

Recent work by Oculus Info has resulted in the development of a tool called nSpace – a “sandbox for analysis” where analysts can assemble and organize information in both formal and ad-hoc structures [13]. In nSpace, the post-it note is used as a flexible metaphor for developing and connecting analysis artifacts. The interface features a large blank canvas on which notes, graphics, and other information can be flexibly arranged. Information can include evidence taken directly from other tools, or it can be added in the form of questions or

concepts recorded by the analyst – allowing combinations of direct and derived evidence.

Oculus has also recently developed a storytelling extension for GeoTime, their space-time visualization environment [7]. With their tools, analysts can take snapshots of the visualization and associate these snapshots with text descriptions that contextualize their findings in a way that is understandable to decision makers. Links to the snapshots from the storyboard allow users to return to the visualization to explore referenced patterns at any time, eliminating the typical separation of visual analytic environments and the tools used to add meaning to their results.

Eccles et al. [7] situate support for storytelling in GeoTime as part of a process to bridge the gap between patterns derived from data and presentable narratives. They describe current systems as good at performing the former task and weak at supporting the latter. It is possible to see similarities in this framework to earlier theoretical processes described by DiBiase [1] and MacEachren [2]. In both cases specialized synthesis tools are called for to transition from analysis to presentation.

The Information Interfaces Group at Georgia Tech has recently developed a text report visualization toolkit called Jigsaw [10]. This toolkit features multiple, coordinated views designed to explore entities derived from large collections of text reports. It also features coordination with Microsoft's OneNote annotation software. Jigsaw users can write down observations and take snapshots of the Jigsaw interface and organize these inside OneNote while working with the visualization. It is unique among recent synthesis support tools in that it uses a tablet interface to facilitate synthesis, enabling users to develop personalized knowledge representations.

The National Visualization and Analytics Center (NVAC) at Pacific Northwest National Laboratory (PNNL) has recently presented a synthesis tool called the Scalable Reasoning System (SRS). The SRS is a web-based environment for organizing bits of information represented as individual post-it notes on a flexible canvas. Users can conduct web searches and use web services to retrieve information, and then represent findings as notes or links between notes [8]. Automated methods can be applied to information in SRS to develop simple visualizations for both text and numeric data – opening the door for subject matter experts to integrate visualization into their work without requiring

substantial training investments. Additionally, artifacts and links in SRS can be given confidence and quality ratings using interactive sliders.

Researchers at the Massachusetts Institute of Technology have introduced a synthesis support environment called the Electronic Card Wall, or EWall for short [9]. The EWall is designed to support collaborators working at a distance as they search for, retrieve, and share information artifacts with each other in a dynamic workspace. Information artifacts are represented in the form of cards that can be arranged on a blank workspace. EWall users can assemble information in personal workspaces and share pieces of their individual workspaces on a global, collaborative workspace.

Aside from tool development efforts, there have been several important contributions to the study of information synthesis through related empirical and theoretical work. On the theoretical side, Heer and Agrawala [14] have developed a design framework for supporting collaborative visual analytics that includes considerations for artifact sharing and histories. In terms of empirical work, Isenberg et al. [15] conducted a study to explore the strategies that teams use when deriving meaning from paper artifacts that include simple visualizations. A complementary study by Robinson [16] characterized how teams develop hypotheses from paper artifacts that represent information from a range of sources.

Each of these projects is working toward the goal of supporting analysts as they synthesize information from a variety of sources in a wide range of analysis situations. The research reported here focuses on understanding the needs of analyst end-users to inform the design of these types of visual analytics synthesis tools to ensure they advance the current state-of-the-art.

4. Interviews

The following sections describe interviews we conducted with analysts to characterize the current methods and tools for conducting synthesis and to elicit ideas for new synthesis support tools. We discuss our participants, the interview format, and the interview questions.

4.1. Participants

Eight disease surveillance and biological/chemical threat analysts were recruited from Pacific Northwest National Laboratory to participate in interviews and individual synthesis experiments (the results of which are currently in preparation for publication). These analysts were recruited with assistance from NVAC scientific staff at PNNL who were asked to identify disease and biological/chemical analysts who would be likely end-users of visual analytics tools currently in development. Each analyst was provided with a stipend by NVAC to compensate for their time spent participating in the research.

Four male and four female analysts participated in this study, working in various roles as analysts at PNNL. Three are research program managers, focused on biological and chemical monitoring and security issues. The remaining five are research scientists, three working in biology and medicine, and two working in chemistry. All of these participants are actively engaged in providing analytical products for clients in a wide array of U.S. government agencies. Projects are typically initiated by client agencies, who issue contracts to PNNL analysts for reports or other analysis products. Participants in this study stated they were currently working on specific topics like avian influenza surveillance, disease modeling, and chemical and biological weapons proliferation.

The problems that analysts in this study undertake are dynamic and intricate – problems that typically involve collaboration among many analysts, each of which has a particular area of expertise. The products they generate are often reports for decision makers, which link together individual analyses into a coherent story. The nature of their work makes these participants likely future users of synthesis support tools.

4.2. Interview format

In July of 2007 analysts recruited at PNNL were interviewed with the goals of eliciting knowledge about the current state of synthesis support tools, and to develop possible design directions for future synthesis support tools.

An hour long structured interview was developed to shed light on how analysts currently conduct synthesis as well as how they envision future synthesis support tools. A structured interview format was chosen to ensure that all participants received the same questions and that answers across the group could be more directly comparable. Structured interviews require all questions to be preselected, placed in a predetermined order, and asked without modification or adlib follow-ups [17]. They do not provide for flexibility like semi-structured or unstructured interviews where the questions can be created or modified during the interview to probe potentially interesting avenues. They do afford readily comparable answers and help to alleviate problems with interviewer bias that can occur when questions are open to modification.

In this research, interviews were intended to collect basic and comparable knowledge about the character of current and projected information synthesis, making a structured format more appropriate than an open-ended approach.

4.3. Interview questions

To explore how synthesis is currently conducted and how it is envisioned in the future, we developed interview questions focused on these research questions:

- What are typical analysis artifacts?

- How do analysts organize results?
- How do analysts develop the “big picture” from their results?
- How do analysts recall prior work?
- How do analysts collaborate with results?
- How do analysts explain their results to others?
- How do analysts handle information provenance issues?
- How do analysts envision future synthesis support tools?

The complete set of detailed interview questions is available online as a supplement to this paper (<http://www.personal.psu.edu/acr181/IV09.pdf>). Figure 3 shows how each topic area corresponds to specific interview questions.

Interview data reported here comes from digital audio recordings that have been transcribed using Transana qualitative analysis software [18]. The transcripts of all eight interviews total 22,057 words. Answers for each question were compiled and evaluated as a group to derive the results reported in subsequent sections. The goals for interview data analysis were to identify areas of agreement and disagreement between analysts.

Topic Areas	Interview Questions
Participant Background	1a,1b,1c,1d
Analysis Artifacts	1e, 2
Organizing Results	3
Developing the “Big Picture”	4, 13
Recalling Prior Work	6
Collaborating With Results	7, 8
Explaining Results	11, 12
Result Provenance	10
New Synthesis Support Tools	5, 7, 10

Figure 3 Topic areas and interview questions

5. Results

For each of the major question topics, interview participants were asked to follow-up their answers with suggestions about how they would envision improving the current state of the art. The following sections cover responses on each topic, and a final section summarizes analysts’ ideas for future synthesis support tools that would improve upon current systems. Where appropriate, some sections include representative quotes from interview participants to support our conclusions. To protect the identity of participants, they are referred to using the generic prefix “PNNL” and a sequential number suffix (e.g. PNNL 1, PNNL 2, and so on).

5.1. Analysis artifacts

A key area of interest for supporting synthesis with visual analytics tools is to develop an understanding of what types of artifacts are most common. Toward that

goal, participants were asked to describe the typical analytical results that emerge from their work. Responses indicate that artifacts can take many forms, including tables, graphs, images, schematics, and text reports. The program managers in this study indicated that text reports were the most common results they worked with, while research scientists mentioned that other types of artifacts were more common.

When asked to describe how artifacts are stored, participants indicated that operating system file folders, Excel spreadsheets, and PowerPoint presentations were typical mechanisms for collecting artifacts. Participants also mentioned that email archives are important as they often contain results in the form of attachments from colleagues. According to participants in this study, shared network storage is the most common method of collaborative artifact storage, and in some cases email is used to share artifacts among multiple users.

5.2. Organizing results

Multiple questions in the interview were designed to elicit the current process of synthesis as it is conducted at PNNL. These questions began with a request for analysts to describe the strategies they use to organize their results. Participants responded that they would typically organize results by projects or topics. PNNL 2 stated that they would develop “piles” of results:

I generally tend to assemble things into an administrative pile, which you know I can ignore...except for that “are we running out of money?” sort of deal. Then I create a reference pile, if you will, and then one for the project or one for the technical work specifically, and then normally kind of a reporting pile. And so things will generally flow out of the reference into the project, then into the reporting one.

When asked if these strategies change depending on the type of project they are completing, analysts indicated (with one exception) that they tended toward a single common strategy for organizing results. Analysts cited the need for consistent reporting and communication with clients as reasons for using the same strategy regardless of the context. PNNL 8 differed from the others, stating that different project goals typically require different strategies for handling results, as some projects they work on deal with only a few key pieces of information, while others may involve many more.

5.3. Developing the “big picture”

Participants were asked to describe methods and tools they use to evaluate the “bigger picture” that situates their results, a portion of synthesis that shifts simple collections of results into coherent information that can be used for reporting. This question revealed some differences between how those who work on disease and biological threats approach this task versus those who work primarily with chemical threats. The

former group cited their domain expertise as the main mechanism for this portion of synthesis. PNNL 5, an infectious disease expert described how they situate their results:

I think I have some kind of ability to see where things are going, you know? And I think I draw a lot of those thoughts from, you know, reading the newspaper, watching CNN, um, I've got a book...on the plague...I mean there's the same patterns that flow, from historical knowledge and other media, giving context.

In contrast, analysts who work on chemical threats indicated that they rely on statistician colleagues to help situate their findings. PNNL 6 describes one such scenario:

The statisticians. Verifying it with someone else, right. But the graphs typically give you the bigger picture to see what is happening over a time period or whatever you're measuring.

In terms of tools for supporting this task, participants stated that Microsoft PowerPoint, Microsoft Word, email clients, and file sharing software are most commonly used. Two participants indicated that in addition to those tools they also use Starlight [19] and IN-SPIRE [20] visualization tools to create data representations that provide overviews of their work.

When it comes time to present the “big picture” in report form, PNNL 6 stated that this task requires the use of outlines:

That's the difficult part, trying to figure out what to include, what's important. Ah, usually we try to create an outline and then fill in the outline as we go, trying to determine how to tell the story in the best possible manner. And I struggle with that. It's not easy.

One participant described how presenting synthesized results requires careful attention to defining the relationships between particular results. PNNL 4 says:

I think that takes a careful analysis of each of those pieces. What they mean, and how they relate, and I think you really have to demonstrate the relation. I think it needs to be very cohesive when it's presented in order to look like it's...a serious result that should be considered as a whole, I think. Because they all sometimes have to depend on each other. I think especially in the business that we're in, we're trying to get to a root cause or a particular type of technical thing...then you need to be careful with how you present the pieces together.

5.4. Recalling prior work

An important aspect of synthesis pertains to the ability for analysts to recall and reuse prior work when that becomes necessary. For example, a risk analysis

describing the economic and social impact of a potential avian flu outbreak might need to be revisited in the wake of an actual outbreak to evaluate the predictions. For the purpose of better situating the current state of synthesis support, analysts were asked to talk about how they approach this problem.

PNNL 7, an analyst who frequently deals with chemical spectra and other associated data mentioned that in addition to laboratory notebooks, automatically encoded metadata is a useful aid to analytic recall:

Oh I guess that's sort of changed over time but, you have a lab book or sort of like an instrument book. And so you've written down something, maybe what you did that day. And a lot of the instruments these days, depends on the instrument, they'll store a lot of the metadata. Store all of the conditions and that sort of stuff. So you can go back and see what conditions you actually used.

PNNL 3 stated that recalling prior work is a non-trivial task, but that it can be aided by keeping track of items by date:

I would say that's a difficult game. Because a lot of times what I find here is that I'll work on a project for say 3 or 4 months, then we're pretty much done...and I will work on other work for 3 or 4 months and somebody will bring me back and say, "Oh we need to do more work in this area." So I've gotten pretty good at organizing where I put stuff and I try to organize it by date and I'm trying to keep track of all the emails that came to me by date so that I can actually track back through and say, "OK, so we were here when I quit."

Several participants indicated that this task is approached most often as a mental exercise, suggesting that tools do not currently play a significant role in aiding this type of synthesis-related task. PNNL 1 described their method for recalling prior analyses this way:

Retracing your steps back to those places, you find the file if it's electronic or paper copy, or it's the papers on your desk...you're trying to find where you were. When you're there you're trying to retrace your steps to catch up and think about where you were. Suppose I built a simulation of a hospital, even a simple one, I'd have to remember what my thoughts were at the time.

5.5. Collaborating with results

Typically analysts at PNNL work on problems that require the work of multiple experts, including collaborators who may be working from multiple locations. To characterize how collaboration impacts synthesis, participants were asked to describe both the tools and methods they use to handle collaboration when it involves managing collections of results.

In terms of tools for managing results among collaborators, participants indicated that they were using

the same tools they used for their personal collections. Microsoft Word and PowerPoint files, Excel spreadsheets, shared network folders, and email are all commonly used for collecting and distributing results among collaborators. PNNL 5 specified how these tools are used in combination, with email as the means for distributing files created in Excel, PowerPoint, and Word:

I have to tell you, the best thing I really have for sharing results, both here but even you know if you're working globally, you know if you're telling your friend at CDC... is really just email. Just email. Here's the email, here's the bottom line, here's the enclosure.

With respect to more general collaborative methods, participants stated that conference calls, regularly scheduled in-person meetings, and email conversations are typically used to manage collaborative synthesis. Once a collaborative project has been coordinated, reports are often developed asynchronously using shared documents and file resources, as PNNL 2 describes here:

There's a couple things. One is that we all collect data and put it in a repository but then we also have a shared synthesis activity where again you think about the page...we all are on travel at different times and whatnot and sometimes it's hard to coordinate and have us all seated in the same place...but if we all have access to the document we can be updating it and adding to it and editing it and so you know there's kind of a timeless collaboration that can occur on the actual product.

5.6. Explaining results

Because synthesis is a transitional stage between analysis and presentation, it is important to examine how results, once synthesized, are moved into the realm of presentation. To explore this portion of synthesis, analysts were asked to describe situations in which results were particularly easy to communicate, and how they were communicated. They were also asked to describe a time when results were particularly difficult to communicate, and how they overcame that problem.

Analysts consistently reported that the easiest results to communicate to their clients were those that the client had expected to see. And the most difficult results to explain are those that the client did not anticipate. Participants indicated that they were usually aware how their work would be perceived (either expected or unexpected) before developing their final report, and that their approach to the report would differ depending on this knowledge. PNNL 4 described the two different scenarios this way:

Yeah I would say if it's an expected result. That's probably the easiest, when they're unexpected is probably when it's the most difficult and then you have to properly surround your results with the supporting information. Go back to references and things.

Similarly, PNNL 2 explained what they have found to be the key factor involved in explaining results to clients:

It's actually kind of more the expectation. He expected that we were gonna come back and reinforce what he thought, it was the fact that we were not reinforcing it that made it difficult, that made it more important that we had data to back it up.

One participant mentioned an alternative scenario in which the results are not easily interpreted by non-experts, making the issue less about whether or not the client agrees with the answer so much as how well the client can understand the science itself. PNNL 7 talked about how this situation occurs:

You know it always depends on who you're trying to talk to. It depends on their background and what they understand and, I mean that's what I've always noticed nowadays because usually the client has a very different background than you, and... what keeps you excited about doing work is going to be very different than what keeps the client excited about doing work.

Software tools were not mentioned by participants in their responses on the issue of explaining results, perhaps indicating that this aspect of synthesis is not dependent on tools as much as it is on the content itself. It shows that there may be a place here for tools to support this task, perhaps by allowing users the ability to quickly attach metadata, credibility measures, or other information to results so that clients can drill down into reports when necessary.

5.7. Result provenance

A crucial aspect of supporting synthesis is to ensure that results can be linked back to the visual analytics software tools responsible for creating them. When asked to describe how this problem is currently handled, all participants stated that current tools do not typically allow them to easily determine who, how, or when results were generated. This is a particularly difficult issue considering that analysts tend to collect results in office productivity software – environments that do not normally recognize analytical results in a way that allows their metadata to be maintained (if they have metadata at all). As PNNL 3 mentions, accessing information that describes result provenance is particularly problematic in situations where analysts are returning to previous work:

We typically run into that a lot, especially from what we did years ago, we went back and looked at some stuff and were like, "how did that get from point a to point b?" You can see it in the write up we did, but the write up isn't as detailed as what happened. So we can probably guess how we got there but...

Participants stated that a promising avenue for future software development would be to support better linkages between results and the tools (and analysts) that generate them. PNNL 4 supported this direction with the following statement:

You know I really don't think I have much in the way of tools to try to maintain that link. I may make a note somewhere possibly if it's...not recorded somewhere in the electronic file or the email traffic. You know "so and so did this and it was based on so and so's work." But ...a lot of times I don't even receive that information I think. There is a lot of value in going back to the originator of the work...who actually had the original thoughts and the original process of developing something, a method or a tool or something.

In general, interview responses indicate that analysts are aware and uncomfortable with the reality that their results are often distilled into small portions of reports for decision makers. When this happens, information about who developed a particular result or how it was developed is rarely included, and current tools do not support automating this task.

5.8. New tools for synthesis support

To suggest directions for future synthesis support tools, participants were asked to identify changes they would make to what they currently use. They provided several different suggestions for future tools, including easy-to-use visualization environments, keyword searchable databases, and synchronized file sharing tools that do not require substantial training. PNNL 1 described one hypothetical solution for quickly locating results in a collection:

When I want to find this file in the file structure I've got to click and click, and why can't I just say I want this file and there's some artificial intelligence or something that can think like you do and "here it is" instead of me, going to click and scroll down to this folder on my email, and from that down. I think this whole time, "Why can't I at least even say what I'm clicking for?"

One participant indicated that current visualization tools are useful for collecting results and ultimately presenting them to clients, but that they are currently too awkward to support regular use, or use by analysts who cannot dedicate substantial time to training:

Well, IN-SPIRE is pretty easy to use, Starlight is adequate from an output perspective, but it is a huge learning curve when you learn how to operate it. In a perfect world I would make it so that it incorporated data much easier than it does now and allows you to structure it without having to...go in and structure the data so that Starlight would actually read it...I mean it was not an easy process. So there's sort of a disconnect, you have to have somebody who's sort of an expert at

putting it in, but they don't necessarily know what you want, so...

When it comes to collaborative synthesis activities, most participants indicated that using email can introduce problems. Email conversations tend to branch into multiple threads, file attachments are not easily managed, and it is unreasonable to expect analysts to spend time trying to manually organize those things, as PNNL 4 states here:

Email is just kind of a mish-mash of a bunch of different things and I just don't have the time to sit down and file through it and try to say "this is on disease, this is on bioterrorism," and have my separate folders.

PNNL 4 envisioned a new software environment that would help coordinate collaborative synthesis in terms of result management as well as project communications:

You've got your communication side of it, and you've also got...are we all working on the same paper...where are we with that? Maybe a tool where you would go in and it records essentially what you have done in this session, and then in an adjoining folder it also has a log of our communications. And also to maybe somehow link in to email for people who are not as technically adept, you know it automatically synchronizes things.

Interview participants envision new synthesis support tools that make it easy to retrieve important information, and that coordinate that information with related project communications. Additionally, they require that these tools are immediately usable given analysts current technology skills. Document sharing and visualization tools are already in use, but their applications are limited as long as usability remains a barrier, and as long as they do not connect in meaningful ways to project communication streams.

Conclusions

Interview responses from analysts at PNNL indicate that synthesis is currently supported through the use of office productivity software and shared network resources. Office productivity tools like Microsoft Word and PowerPoint are commonly used to organize and share results, both in individual and collaborative settings. Visualization tools are used by some to develop synthesized overviews of information for presentation to decision makers. Email is often used as a method of moving these resources around among multiple people. In general, current tools and methods do not adequately support analysts when they need to revisit and/or revise past analyses.

Client expectations are the deciding factor for the degree to which results are particularly easy or particularly difficult to explain. When analysts know their results will conflict with a client's expectations,

they develop reports that saturate their findings in corroborating details. Current tools do not help analysts easily develop reports with varying levels of detail to match these circumstances.

Participants envision synthesis tools that coordinate project communications with the results they have generated. This would also help alleviate problems noted with determining/maintaining result provenance information. New tools should be usable with minimal training and assumed technological expertise, so as not to impede work progress. Visualization tools are promising candidates for further development in this direction, as their outputs are generally easy for decision makers to interpret.

The results of our interviews with technical and bio/chemical security analysts at PNNL reveal the methods analysts currently use for information synthesis. Analysts also suggest a variety of ways for improving upon current methods with future systems. In general, analysts are currently making use of a wide range of tools that were not originally designed to collect, organize, and add meaning to analytical results.

While our work reflects the views and experiences of a small number of analysts from a single organization, such experts are difficult to recruit in large numbers from a range of organizations. Future studies of additional analysts with different backgrounds and from different organizations would certainly help to test our conclusions. Also, it seems particularly important to conduct observational studies of analysts actually performing the tasks they described in interviews, and this will also augment recent empirical studies of synthesis in laboratory settings [15, 16]. Such studies can reveal additional important details about the tools and methods that analysts use when synthesizing information.

Our findings provide important background information and context for developers of future visual analytics tools that aim to support information synthesis. A wide array of visual analytics tools are in development to support information synthesis tasks, and the research we report here is one of the first examples of a needs assessment study intended to shape the design of such tools based on analyst feedback.

Acknowledgements

This research is supported by the National Visualization and Analytics Center, a U.S. Department of Homeland Security program operated by the Pacific Northwest National Laboratory (PNNL). PNNL is a U.S. Department of Energy Office of Science laboratory.

References

[1] DiBiase, D., *Visualization in the Earth Sciences*. Earth and Mineral Sciences, Bulletin of the College of Earth and Mineral Sciences, The Pennsylvania State University, 1990. **59**(2): p. 13-18.

[2] MacEachren, A.M., *How Maps Work: Representation, Visualization and Design*. 1995, New York: Guilford Press. 513.

[3] Gahegan, M., *Beyond tools: visual support for the entire process of GIScience*, in *Exploring Geovisualization*, J. Dykes, A.M. MacEachren, and M.J. Kraak, Editors. 2005, Elsevier: London, UK. p. 83-99.

[4] Andrienko, G., et al., *Geovisual analytics for spatial decision support: setting the research agenda*. *International Journal of Geographical Information Science*, 2007. **21**(8): p. 839-857.

[5] i2, *Analyst's Notebook*. 2007, i2 Incorporated: McLean, VA.

[6] Wright, W., et al. *Advances in n-space - the sandbox for analysis*. in *International Conference on Intelligence Analysis*. 2005. McLean, VA.

[7] Eccles, R., et al., *Stories in GeoTime*. *Information Visualization*, 2008. **7**(1): p. 3-17.

[8] Pike, W.A., et al. *The scalable reasoning system: lightweight visualization for distributed analytics*. in *IEEE Symposium on Visual Analytics Science and Technology*. 2008. Columbus, OH.

[9] Keel, P.E., *EWall: a visual analytics environment for collaborative sense-making*. *Information Visualization*, 2007. **6**(1): p. 48-63.

[10] Stasko, J., et al. *Jigsaw: supporting investigative analysis through interactive visualization*. in *IEEE Symposium on Visual Analytics Science and Technology (VAST 2007)*. 2007. Sacramento, CA.

[11] Gershon, N. and W. Page, *What storytelling can do for information visualization*. *Communications of the ACM*, 2001. **44**(8): p. 31-37.

[12] Xu, J. and H. Chen, *Criminal network analysis and visualization*. *Communications of the ACM*, 2005. **48**(6): p. 100-107.

[13] Wright, W., et al. *The sandbox for analysis - concepts and methods*. in *Conference on Human Factors in Computing Systems (CHI 2006)*. 2006. Montreal, Canada: Association for Computing Machinery.

[14] Heer, J. and M. Agrawala, *Design considerations for collaborative visual analytics*. *Information Visualization*, 2008. **7**: p. 49-62.

[15] Isenberg, P., A. Tang, and S. Carpendale. *An exploratory study of visual information analysis*. in *ACM Conference on Human Factors in Computing Systems (CHI 2008)*. 2008. Florence, Italy.

[16] Robinson, A.C. *Collaborative synthesis of visual analytic results*. in *IEEE Symposium on Visual Analytics Science and Technology*. 2008. Columbus, OH.

[17] Silverman, D., *Doing qualitative research*. 2nd ed. 2004, London, UK: Sage Publications, Ltd. 416.

[18] Woods, D.K. and C. Fassnacht, *Transana*. 2007, Wisconsin Center for Education Research: Madison, WI.

[19] Risch, J.S., et al. *The STARLIGHT information visualization system*. in *First International Conference on Information Visualization*. 1997. London, England: IEEE Computer Society.

[20] Wong, P.C., et al. *IN-SPIRE InfoVis 2004 Contest Entry*. in *IEEE Symposium on Information Visualization (INFOVIS 2004)*. 2004. Austin, TX: IEEE Computer Society.