



Evaluating System Sufficiency in a Multimodal, Multiuser Sensemaking Environment Designed for Intelligence Analysis

Shannon Briggs^(✉)

Rensselaer Polytechnic Institute, Troy, NY 12180, USA
briggs3@rpi.edu

Abstract. This paper discusses a prototype brainstorming tool designed for sensemaking in intelligence analysis and integrates multimodal input in an immersive environment that supports multiple users. This digital tool has been designed using structured analytic techniques already in use through analog methods in the intelligence analysis domain. We discuss an ongoing user study that presents preliminary findings for tool, and suggestions for future development.

Keywords: HCI methods and theories · Multimodal interface · Evaluation methods and techniques · Qualitative and quantitative measurement and evaluation · User experience

1 Introduction

There have been many digital tools developed for defense and national security venues as a way to enable intelligence analysts and other defense workers to forage for and make sense of information. However, of the software that has been developed, comparatively few overall have been reliably adopted by the intended end users. We propose that developing software that is developed through scenario-based design and sensemaking processes informed by intelligence analysts will create a final product that will be cognitively more approachable and therefore more useful and usable. Currently, the digital brainstorming tool developed and implemented in the Cognitive Immersive Systems Lab (CISL), has been informed by the sensemaking process as described by Pirolli and Card's model [1]. The digital brainstorming tool described in this paper has been designed to support the information-foraging process. We are attempting to transform the analog structured analytic brainstorming technique into an immersive multimodal. We have implemented verbal and gestural technology into the digital, and we have enabled a personal view and a group, or global, view. In this paper, we present preliminary findings of a user study that is currently in progress of the prototype brainstorming digital tool.

1.1 Previous Work

The following section discusses software that has been developed for the intelligence analysis domain, as well as specific features that have been implemented in these tools to enable sensemaking. These works establish a history of development for the range of functions enabled in software for the intelligence domain.

Taxonomy for Tools Developed for Intelligence Analysis. Prior tools developed for intelligence analysis has approached two of the major problems that analysts face in completing their work: sifting through and filtering large amounts of text data for relevance and representing and manipulating data for improved sensemaking. The former group of tools focus on information management tools, that allows analysts to conduct quantitative measures on text corpora, and provide analysts a space to formalize insight generated by their own expertise and aided by the information provided by the software. The latter group of tools focuses improved sensemaking allowing analysts to represent the data in ways that more easily allow them to draw insights, such as generating timelines, representing data as concept bubbles, and so forth. In both groups the effort is aimed at improving sensemaking for intelligence analysts, while also reducing the cognitive burden of searching for, filtering, and classifying information.

Information Management Tools. Information management is of the major portions of the information-foraging process, and includes searching, filtering, sorting, and marshalling evidence, as described by Pirolli and Card [1]. This stage is important for analysts as they begin to find information that is important to the issues they are concerned with. The cognitive challenge inherent in this stage is information overload, and digital tools, such as those described below, are aimed at assisting analysts in managing information volume, and using algorithms to assist in searching.

Jigsaw is software designed for intelligence analysis especially through text analysis. Gorg et al. [2] discuss the capabilities of the software in their article. Jigsaw is designed to support the sensemaking activities surrounding collecting and organizing textual information by intelligence analysts. The Jigsaw system is designed to provide visualization for different perspectives on information in documents and it supports evidence marshalling through a Shoe Box view. The earliest version of the software focused heavily on visual representation of relationships between entities but did not provide any kind of text analysis. One of the major findings of creating this software was that software functions cannot replace the reading of reports. Repeated careful reading of selected texts tended to be the preferred method to understand the information in texts. As a result, the Jigsaw system incorporated the ability to summarize and cluster similar important text information. From there, the software used packages such as GATE, LingPipe, the OpenCalais system, and the Illinois Named Entity Trigger, to import data from documents.

DIGEST is another tool designed for intelligence analysis, described by Benjamin et al. [3]. Its main capabilities are extracting data from text, such as sentiments, social influence, and information flow structures; the tool also has exploratory data analytics, and finally it uses the stored results to create various knowledge products. After the data collection and processing stage, where analysts can configure the tool to collect data on specific topics, the tool develops a template for information reporting. Finally, the analyst can populate the template with the information the tool has collected. The analyst can choose what information they want included, as well as add any of their own insights to the product.

E-Wall is a visual analytic environment design to support remote sensemaking through the development of what the authors call a virtual transactive memory [4]. The software focuses on object focused thinking, where information is represented as an object, and users construct semantic relations between them. Ewall uses computational agents to infer relationships among data and offer customized insights to users. The E-wall layout is designed to allow users to collaborate while working on information and to allow users to manipulate data in object-like chunks. The E-Wall uses two computational agents to manage information flow, and infers relationships among data types, and another that evaluates databases and suggests data to the user. The E-Wall allows users to navigate large amounts of data independently and minimizes the need for verbal interaction.

Sensemaking Management Tools. Sensemaking is in the second part of the sense-making loop described by Pirolli and Card [1], and includes integrating new information into established schema, generating hypotheses, and presenting analysts' judgements and interpretations. This second group of software can be considered sensemaking management tools, as their scope is to better incorporate information into different types of cognitive frameworks, such as new visualizations, concept maps, and timelines.

The Polestar intelligence analysis toolkit is one of the earliest software suites designed for intelligence analysis [5]. Polestar includes a snippet view of texts, where users could highlight and drag text to the portfolio view for later analysis, and records metadata about the text. Polestar also included a way to start knowledge structuring, such as a wall of facts similar to the sticky-note exercise taught in intelligence analysis classes. This software included a timeline feature, to allow analysts a way to visualize relationships in data. Polestar included an argument tree editor, allowing analysts to structure and formulate hypotheses in a visual fashion. The dependency viewer allowed users to trace back where a document or object was found in the dependency network.

VisPorter, is a collaborative text analytics tool aimed toward allowing sensemaking in a collaborative environment [6]. The software is meant for multi-user engagement and the designers focused on different elements such as haptic touch, lighting, and to explore how people forage for information to share hypotheses. The VisPorter software includes the Foraging tool, which contains the document viewer and the concept map viewer, and the Synthesis tool, which allow users to share information found individually with the foraging tools. Some of the features included in this software was gesture-based interaction, with an example of someone with a small display flipping a document off the left side of their device, and having it be shared and dropped on the right side of a synced large display.

Wright et al. [7] introduce Sandbox, which explains human information interaction capabilities, such as ‘put this there’ cognition, automatic process model templates, gestures for fluid expression of thought, assertions with evidence, and scalability metrics. The authors take note of the use of Post-its by to organize and sort ideas, and is then translated into a feature called MindManager, which employed concept map strategies to allow diagrammatic visual representations. The software is designed to adapt to different types of analysts and analytical styles. The software also incorporates a source attribution and context function.

GeoTime is a software system designed for analysts to map geo-temporal events. Geo-Time explores part of intelligence analyst’s process of using narration to make sense of events of interest [8]. The software employs a space-time pattern finding system, which relieves the analyst from effort of searching for common patterns and events. The second part of the system relies on visual annotations, which takes the visual information and attaches relevant information. The final part of the software is a text editor that allows analysts to make relevant comments on the found information. GeoTime uses a collaborative environment but also emphasizes a data-aware object, where annotations are embedded in time and space, so these become a new piece of information connected to the found information. GeoTime is also interested in allowing analysts to work on a meso-level, such as behavioral trends, events, and plots, rather than an individual unit.

1.2 Digital Brainstorming Tool

The digital brainstorming tool developed by CISL allows a digital space for human analysts to put into practice cognitive techniques already familiar to them. The digital brainstorming tool can be considered as an information visualization tool as well as a schema representation tool. The tool was designed to enable human collaboration and sensemaking during intelligence analysis, and we hope to achieve create a cognitively more accessible product by applying the intended users’ domain knowledge to our software development.

Our approach to designing a tool that will assist in sensemaking for the domain was to consult and integrate techniques described in educational materials into the software design of the brainstorming tool. The specific process for the digital tool is a structured analytic technique called brainstorming, and analysts write salient pieces of data on sticky notes and creates topic groups from these notes. This structured analytic technique incorporates timed periods of reading, analysis, and collaboration, to gain a diverse perspective about the information. Our research anticipates that communication between users is an important part of the collaborative process and therefore our tool integrates deliberate periods of interaction among users to collaborate on theories.

One of the major insights from discussion with analysts is that digital tools that are too cumbersome to use, or tools that require analysts to step outside of their workflow to interact with tend to be left behind.

The digital brainstorming tool is a digital tool that is based on the pen and paper version of the structured analytic technique used by intelligence analysts, as described by Beebe and Pherson [9]. The digital tool is displayed via five projectors onto a 360-degree panoramic screen. The digital brainstorming tool allows viewers to generate digital sticky notes, and users can interact with the system through gesture technology using Kinect cameras to capture body frame and gestures [10], and through verbal commands to the Watson system using lapel mics. This system allows user to interact with the digital tool in a manner that mimics the analog interactions that are found in the structured analytic tool, and provides a more immersive experience. The digital brainstorming tool has two major components, the global view and the personal view. The global view can be seen in Fig. 1, and the personal view can be seen in Fig. 2. The global view is projected onto the panoramic screens in the immersive environment, allowing users to discuss with each other and interact with the system; the personal view is accessed through any personal device that is equipped with a web browser.



Fig. 1. Global view of the digital brainstorming tool.

The system is accessible separately and simultaneously through the global and personal views, allowing a myriad of interaction types for users. Users are able to interact remotely with the personal view, viewing others' notes on the global view, and adding and categorizing their own and others' through the touch and type interface of technology such as tablets and laptops. The global view allows users to move sticky

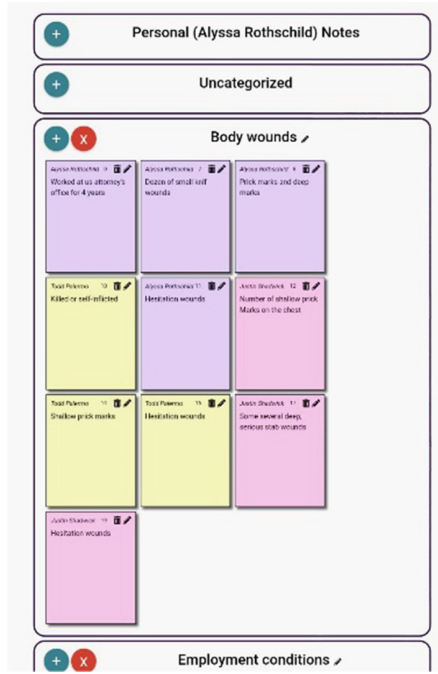


Fig. 2. Personal view of the digital brainstorming tool.

notes on the shared 360-degree screen, and issue verbal commands to interact with the system (Table 1). The breakdown of the current capabilities of the system are as follows:

Table 1. Capabilities of the digital brainstorming tool

Functions	Global view	Personal view
Create note	Voice	All functions are enabled on the personal view's UI
Delete note	Voice	
Move note	Gesture/voice	
Edit note	Voice	
Move note	Gesture/voice	
Categorize note	Gesture/voice	
Create category	Voice	
Delete category	Voice	
Rename category	Voice	

2 User Study Methodology

The user study was designed with participatory design and sensemaking design informing our approach. The study is an A/B study, with participants engaging in an analog brainstorming exercise as outlined by Beebe and Pherson [9]. The study was conducted in 1-hour sessions, with small groups of two to three participants engaging in 30 min segments of first the analog tool and then 30 min of the digital tool. Participants completed the brainstorming session using the Jonathan Luna case study, which is in wide use in training materials developed for intelligence analysts [9]. Participants were given different text segments of the case for the analog portion and the digital portion, with two text segments used overall. We chose the Luna murder case study as we felt it would be more approachable to a wider participant base and required less prior knowledge or interest in the intelligence analysis domain. We also use the analog session of the study to familiarize users with the format of the structured analytic technique outlined by Beebe and Pherson [9], as it is a specific technique that is unlikely to be taught outside of intelligence analysis courses. While later-stage user testing of the tool will involve intelligence analysts to provide their feedback and input for the tool, we were initially more interested in user input regarding basic usability functions that would not require particular expertise in any field or discipline. Particularly, we wanted to explore the validity of our hypothesis: that the digital brainstorming tool would be functional as well as the analog brainstorming tool as a sensemaking tool for novice users.

We used user surveys for both analog and digital phases. Participants completed separate user surveys after each 30-minute session and were invited to participate in a brief question and answer session at the end of each 30-minute segment. Our surveys asked questions about previous experience with experience in intelligence analysis, experience with brainstorming exercises and software. Participants were recruited from the student body at the university, and overall were novel to this structured brainstorming techniques. All participants were first time or novice users of the interaction technology with the digital tool. We felt this was important to determine the tool's usability to users who are at baseline unfamiliar with the underlying technology, as we posit that most potential users are not going to be familiar with gesture recognition or voice recognition technology in context of an intelligence analysis sensemaking tool. We used think-aloud protocol for our digital tool segment, inviting participants to voice their thoughts as they interacted with the tool, so that we could capture their unfiltered impression of the digital tool as they worked. A more thorough explanation of the digital and analog portions of the experiment is described below.

We captured video data and audio data for the experiment, and logged user's actions with the system to determine which users made which notes and categories. Video data is used to record the analog sensemaking process, as well as to give a complete picture of the digital sensemaking process. We are also able to understand users' observations and utterances as pertaining to the system and their interactions.

Analog Testing Procedure. Participants were given a text copy of a text segment of the Jonathan Luna murder case, and each participant was given a pack of differently colored sticky notes. Participants were given a 5-minute period to read the provided text segment, then were given another 5 min to write pieces of evidence from the case study on separate sticky notes. Participants then spent 5 min putting their sticky notes on a shared wall, with each participant laying out their own sticky notes in separate small groups according to topic. Participants then spent 5-10 min discussing and rearranging the notes into shared groups, sorted by common ideas into categories. A brainstorming facilitator was used to time each of the five-minute segments and move participants into the next stage of brainstorming.

Digital Testing Procedure. Participants were given a brief tutorial describing the functions of the digital brainstorming tool, and tablets were distributed to each participant to allow them to access the personal view. Participants were then given a text copy of the second segment of the Jonathan Luna case study, which was also available on the global view of the system and was viewable by all participants. Participants were given 5 min to read the text segment, after which they were given a further 5 min to write pieces of evidence using the personal view on the tablets. Participants were then given 5 min to then use the tablets to send their notes to the global view and sort their notes into separate groups using the gesture technology. Participants were finally given a final 5 min to discuss their notes with each other and rearrange notes and create categories using the personal view and the global view. As with the analog exercise, a brainstorming facilitator timed the five-minute segments and moved participants along the brainstorming session.

Hypotheses. As our user study is an AB comparison study, we are interested in how the digital brainstorming tool performs compared against the traditional brainstorming pen and paper brainstorming tool. We are interested to see if the digital brainstorming tool is as sufficient as the pen and paper process in allowing users to make sense of information as they are working through the case study. We decided to focus on two major aspects of the brainstorming tool that are reflective of the brainstorming process regardless of format: the overall number of sticky notes created per session, and how many notes were put into categories. These two features emphasize participants' interaction with the brainstorming tool and interaction with other participants, as number of sticky notes is the result of participants examining the sample case study, extracting of information from the case study, recording data, and then finally discussing information that was found individually as a group, and notes grouped per category allows us to see how participants' used the system to agree on how the information can be categorized by consensus,

H1: The digital brainstorming tool is as sufficient as the traditional analog brainstorming tool in enabling sensemaking if users are able to create as many notes within the digital system as they can with the analog tool.

H2: The digital tool is sufficient as a sensemaking tool compared to the analog tool if users are able to categorize notes as efficiently as they do with the analog tool.

3 Results

3.1 Digital Sticky Notes

As this digital tool is based on an existing analog cognitive exercise, our current interest is to discover if the digital tool is as sufficient as the analog tool; that is, if the digital tool as a medium for participants to conduct a brainstorming session is as easy to use as the established analog tool. Our way of measuring this is number of sticky notes generated per group per session, as we assume that if the digital tool were much harder to use than the analog tool, then participants wouldn't be able to create as many sticky notes as they had in the analog portion. Therefore, for H1 we examined are number of notes users were able to generate per session. Our hypothesis is that the digital tool can be considered as useful as the previously developed analog tool if users were able to develop as many notes or more on the digital tool as they were able to create using analog sticky notes. The table below shows a breakdown of number of notes users generated across 5 sessions between the analog and digital sessions (Table 2).

Table 2. Number of sticky notes across analog and digital sessions

Session number	Analog session	Digital session
Session 1	16	15
Session 2	25	19
Session 3	20	6
Session 4	16	13
Session 5	46	31

Overall, we had largely positive results in comparing the number of notes users were able to create in the digital tool versus the analog tool. We found that users created nearly as many notes with the digital tool as with the analog tool, save for session 3, which was terminated due to technical problems. We do note that while there are fewer notes overall for the digital session which we further examine in the discussion section.

3.2 Digital Categories

The second point we examine is comparing number of categories and number of notes per category between the analog brainstorming tool and the digital brainstorming tool. Creating categories is a major discussion point between participants, so the number of categories can help us understand if participants are able to use the digital tool to make sense of the information they have collected. The number of notes per category helps us understand if participants can interact with the functions of the system to arrange their notes in a way that makes sense to them. The Table below is a breakdown of categories and notes per session (Tables 3 and 4).

Table 3. Analog categories and number of notes in categories per session.

Session number	# Categories	Category 1	Category 2	Category 3	Category 4
Session 1	3	5	10	1	X
Session 2	4	4	7	5	X
Session 3	3	5	3	5	X
Session 4	4	2	4	3	7
Session 5	3	17	15	15	x

Table 4. Digital categories and number of notes per category per session.

Session number	# Categories	Category 1	Category 2	Category 3	Category 4	Uncategorized
Session 1	2	2	3	2	1	1
Session 2	3	3	2	8	X	X
Session 3	3	7	4	2	X	2
Session 4	3	6	5	2	X	X
Session 5	4	10	7	9	5	X

For H2, we found that users generally created one fewer category using the digital tool than the analog tool and had more uncategorized notes than with the analog tool. We posit that this may be due to the somewhat more difficult process of creating categories through the digital tool function. An alternative interpretation is that using the digital tool allows for more discussion time, as we observed it takes less time to share notes and create categories, which we confirmed in session five that that users of the digital tool create more categories when greater time is allowed during the discussion phase. We feel that further modifying the ability to create categories in a more natural fashion may improve users' ability to create categories and properly categorize notes.

3.3 Discussion

While our current participant pool response is limited in size, we are continuing to run user studies to gain further understanding of the ability of the digital brainstorming tool as a sensemaking aid. We found that users' responses to the tool became more favorable as we implemented usability improvements to the software between user testing sessions. Participants gave valuable feedback in both functionality interactions and interfaces features, which we are continuing to implement for further user studies. Overall, the user experience with the digital tool was mixed, but improved with features that focused on ease of use. Based on user responses, we implemented the ability to edit category names, a clearer way of editing notes via the personal view, and a way to view voice commands on the global view. Similarly, most users responded positively to basic aspects of the digitization of the tool, such as freedom offered by the digital format of the sticky notes, being able to read others' notes clearly, being able to read transcripts of the conversation, and voice interactions. In general, participants felt that the gestures system was intuitive.

Sticky Notes. The content for the sticky notes in the digital session are largely similar across sessions, indicating that the tool is sufficient in allowing users to extract a baseline level of information from the sample case study. Based on the results of the user study so far, we find that the digital tool can aid users in creating nearly as many sticky notes and categories as with the pen and paper tools. The reduced number of notes in the digital sessions may be the result of the changed case study text between the analog and digital session rather than an insufficiency in the tool, which we will examine during future user studies. Finally, while the gestures system is able to support multiple users, users often take turns in using the system regardless, as though it were a single user system. We anticipate that with further modifications to the software to ensure stability and robustness that users should be able to surpass number of notes and categories generated, especially considering that the digital system is able to support an unlimited number of notes or categories generated. This feature was particularly important, as it reflects the success of the digital tool in allowing users to create as many notes as necessary to reflect the personal understanding of the sample case study used as needed by each individual user. Given the prototypical nature of our tool, we are encouraged that with modifications for usability and stability, that the digital brainstorming tool shows potential in enabling sensemaking for our use case.

Categories. One of the major differences in the analog and digital tool is the way categories are created and named. As per the pen and paper method, the creation of categories is delineated by notes grouped in proximity, and a label to describe the content of the category is agreed upon by consensus; the recording of that label can take several forms, either above the categories themselves or on a separate paper, and so on. In our study, participants verbally relay category names to the discussion facilitator to discuss their choices for the analog session. However, the digital tool changes that process by manner of the technology; users switch from the personal view, which can be accessed on any web browser, to the gesture and voice technology of the global view. Multiple participants commented on the ability to clearly read both category names and note content, which we believe is an important improvement from the pen and paper tool. We believe that the affordance of clearly written note content helps facilitate discussion of the case study and creation of category labels, as users are not forced to parse others' handwriting, and can more effectively engage in the presented material.

In early iterations of the prototype tool, deletion of a category resulted in deletion of the notes in that category, which also dissuaded users from creating more categories; however, with user feedback, we implemented the ability to transfer notes within a deleted category to the shared view to preserve their content. We feel that further modifying the ability to create categories in a more natural fashion may improve users' ability to create categories and properly categorize notes.

One of our earliest observations in users engaging structured analytic brainstorming with the digital tool was the collapsed time requirement in sending notes up to the shared view. In the analog sessions, the element of bringing physical sticky notes up to a shared wall space is timed, and usually takes up the allotted time for the activity. We have found, however, that in allowing users to instantly send notes to the shared view reduces that time drastically. This allows for a longer discussion period between participants during the final phase where participants discuss note content and how to categorize their notes.

4 Conclusions, Limitations, and Future Work

We have developed a digital sensemaking tool for intelligence analysis based on analog structured analytic techniques, which were created by intelligence analysts. This software differs from previous work in the field, as discussed above, as we implement our software in an immersive environment to mimic analog interactions, and we have designed the tool to be informed by the cognitive techniques developed by intelligence analysts, which we believe will be more useful and usable to the intended user. Our user study was developed using open source information, available through text books and other educational materials. We plan on refining the functions of our system through direct feedback from intelligence analysts who have worked in the domain.

Our digital brainstorming tool is a prototype, and as can be seen through the progress of the user study is being modified to ensure stability, responsiveness, and robustness in order to support the sensemaking process in intelligence analysis. Future modifications of the tool will be focused on implementing more gestures, overall smoothness of interaction with the different multimodal inputs, and ways in which to link the brainstorming tool with other digital sensemaking tools. We will also be introducing the digital tool to participants in a more formalized demo before testing in order to allow future users to acclimate interacting with the tool.

For future work, we will continue to run our user study to add to the body of data we have collected, and we will supplement our analyses of the data gathered with statistical insight. We are particularly interested in understanding how the medium of the tool affects issues such as how many sticky notes are generated and how many categories are created. We are soliciting feedback and advice from former intelligence analysts, and will run a user study to gain their reactions to the system. We also plan to investigate the issue of cognitive load and how our digital tool can reduce elements of cognitive load during the intelligence analysis process.

Acknowledgements. I would like to acknowledge Matthew Peveler, Jaimie Drozdal, Lilit Balagyozan, and Hui Su, who have all contributed to the material presented in this paper.

References

1. Pirolli, P., Card, S.: Sensemaking processes of intelligence analysts and possible leverage points as identified through cognitive task analysis. In: Proceedings of the 2005 International Conference on Intelligence Analysis, McLean, Virginia (2005)
2. Görg, C., Liu, Z., Parekh, N., Singhal, K., Stasko, J.: Jigsaw meets Blue Iguanodon - The VAST 2007 Contest. In: In IEEE Symposium on Visual Analytics Science and Technology (2007)
3. Benjamin, P., Madanagopal, K., Erraguntla, M., Corlette, D.: Distributed Information Gathering, Exploration and Sensemaking Toolkit (DIGEST). In: Proceedings on the International Conference on Artificial Intelligence (ICAI); Athens. pp. 449–455. The Steering Committee of The World Congress in Computer Science, Computer Engineering and Applied Computing (WorldComp), Athens, United States, Athens (2016)

4. Keel, P.E.: EWall: a visual analytics environment for collaborative sense-making. *Inf. Visualization* **6**, 48–63 (2007). <https://doi.org/10.1057/palgrave.ivs.9500142>
5. Pioch, N.J., Everett, J.O.: POLESTAR: collaborative knowledge management and sensemaking tools for intelligence analysts. In: *Proceedings of the 15th ACM International Conference on Information and Knowledge Management*, pp. 513–521. ACM (2006)
6. Chung, H., North, C., Self, J.Z., Chu, S., Quek, F.: VisPorter: facilitating information sharing for collaborative sensemaking on multiple displays. *Pers. Ubiquit. Comput.* **18**, 1169–1186 (2014)
7. Wright, W., Schroh, D., Proulx, P., Skaburskis, A., Cort, B.: The sandbox for analysis: concepts and methods. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 801–810. ACM, New York (2006). <https://doi.org/10.1145/1124772.1124890>
8. Eccles, R., Kapler, T., Harper, R., Wright, W.: Stories in GeoTime. *Inf. Visualization* **7**, 3–17 (2008). <http://dx.doi.org.libproxy.rpi.edu/10.1057/palgrave.ivs.9500173>
9. Beebe, S.M., Pherson, R.H.: *Cases in Intelligence Analysis: Structured Analytic Techniques in Action*. CQ Press, Washington, DC (2014)
10. Divekar, R.R., et al.: CIRA—An Architecture for Building Configurable Immersive Smart-rooms. 10