Translating the Pen and Paper Brainstorming Process into a Cognitive and Immersive System

Matthew Peveler[1(✉)], Shannon Briggs[1], Jaimie Drozdal[1], Lilit Balagyozyan[1], Chang Sun[1], Michael Perrone[2], and Hui Su[2]

[1] Rensselaer Polytechnic Institute, Troy, NY 012180, USA
{pevelm,briggs3,drozdj3,balagl,sunc4}@rpi.edu
[2] IBM Thomas J Watson Research Center, Yorktown Heights, NY 10598, USA
{mpp,huisuibmres}@us.ibm.com

Abstract. Within this work, we introduce a digital sticky note tool to be used by analysts for a brainstorming session within a cognitive and immersive system (CAIS). This tool was setup to mimic and enhance the traditionally pen-and-paper process. By utilizing a digital interface, we allow for more feature-rich note taking including pictures and hyperlinks in addition to just text, as well as an increased amount of captured metadata about each created note. To achieve parity with the non-digital process, our interface allows for both individual usage via personal devices (e.g. tablets, laptops) as well as for coming together as a group and collaborating via a multi-modal interface, supporting commands through a mixture of voice and gestures.

Keywords: Human–centered computing ·
Collaborative and social computing systems and tools, technology ·
Gesture and eye—gaze based interaction, technology ·
Second screen (multi screen scenario)

1 Introduction

Much of contemporary research and design into tools for intelligence analysts focus around the use of tools aimed at individuals. However, much of the analysis is done within a social environment, with analysts working in groups to come to a conclusion about the data under review. While it is still possible to utilize these tools, this often comes at the price of having to email data around so that each analyst can use the tool, or by grouping around a single workstation that is then driven by a single person through the use of keyboard and mouse, or a single touchscreen. However, often times the best solution is to fallback to a more traditional "pen-and-paper" version of the tool such that all members of the group can then participate simultaneous.

One such tool that we examined for this work is sticky notes, and their use in the brainstorming technique for intelligent analysts. The brainstorming

structured analytic technique is a cognitive tool used by analysts to help generate new ideas and opinions about materials relevant to their job. In addition, it is one of the earliest techniques that can be carried out during the analysis process [2]. The brainstorming technique is a multi-step process that involves data collection, analysis, and then group discussion. The purpose of the exercise is to reduce cognitive bias, and to allow a diversity of perspectives to be taken into consideration when dealing with a problem or scenario. For this work, we utilize a multi-step process for a brainstorming session, wherein there are at least two analysts and a discussion facilitator. The process starts with the analysts operating independently to analyze and document the salient details of a scenario onto notes. The analysts then come together as a group, synthesizing a final result of notes as well as creating categories to group them.

In this paper, we present a full description of the pen-and-paper brainstorming process, including a list of the primary high level actions that are taken. We then describe what is a cognitive and immersive system, and how we utilize it to allow and accomplish handling both multi-users and multi-inputs for a group of analysts. In the next section, we describe the digital tool that we've created and how it can be used by analysts. We then present the results of a user study that we conducted comparing and contrasting the two methods. We then conclude the paper with a discussion of the results and promising future lines of work.

1.1   Prior Work

Electronic sticky notes have been proposed before in digital brainstorming literature, and varieties of such a tool have been implemented. Prior research has examined the application of sticky note tools in group and collaborative settings [5], in mediated group work accomplished remotely, and how sticky notes can be used to define affinity groups in collaborative work [10]. Jensen et al. [6] conducted a study comparing the use of traditional analog sticky notes to a digital sticky-notes tool, and concluded that the digital sticky notes were superior in terms of increased note interaction, clustering, and labelling. Existing digital sticky note tools can be found in examples such as Discusys [10], ECO-Pack [9], Padlet, and Quickies [7]. These tools demonstrate the power of moving to the digital medium, in allowing for embedding of digital content (e.g. pictures, videos, hyperlinks) alongside the text, which is not possible in the the analog version. However, these tools focus on individual usage for making and taking sticky notes, and not necessarily a collaborative group interface.

2   The Pen and Paper Brainstorming Process

Traditionally, analysts utilize a "pen-and-paper" process to conduct their brainstorming process (though they may use a whiteboard, or some other tool instead of paper). We focus on the usage of sticky notes through the process. During the initial portion of the brainstorming process, analysts utilize a personal workspace to generate notes that they think are relevant to what they're studying. After

some time spent generating these notes and potential categories for them, the analysts then come together as a group to generate a final working set of notes. At first, the analysts would put up all their ideas together on a board, removing duplicates. The analysts would then discuss what potential categories they think the notes could fall under. After deciding on a set number of categories, the analysts then assign categories to each note. This process can be derived into the following number of high level actions:

– Generate notes in a private workspace
– Share notes in an uncategorized workspace
– Create and delete categories for notes
– Assign notes to categories

These actions are then inline with prior work done on how experts and non-experts approach this task, and the process they go through [1].

3   Cognitive and Immersive System

To support the translation into a digital process, we utilize a cognitive immersive room architecture [4] in building our cognitive and immersive system (CAIS). At its core, we have lapel microphones (one per user) that pick up what participants say to the system and each other. We utilize the IBM Bluemix Speech-to-Text service to then transcribe the speech to a textual representation, and then Watson Assistant to generate an intent and entities for each utterance that contain our "trigger word" for the system. In addition to the microphones, we utilize a Kinect camera to monitor users' actions, primarily focusing on what the users might be pointing at as well as simple (open, closed, and pointing) gestures the user might do with their hand. These intents and gestures are fed into a central executor that then drives an action within the room. The system then has available to it several displays (projectors or screens) as well as speakers.

The system utilizes the Electron framework (in what's called the *display-worker* to show content on the screens, allowing for showing websites that are placed within a grid defined within the available display space, allowing us to show both internally developed websites. The speakers are used to play synthesized voice from the system utilizing the IBM Bluemix Text-to-Speech service. The diagram of components are shown in Fig. 1. The modules that are connected via red lines utilize RabbitMQ to send and receive JSON objects. The blue lines represent communication via HTTP GET/POST methods. The purple lines represent communication through websockets. The yellow lines represent connections to physical hardware.

The components that were developed for this work were the *spatial-context-bridge*, *live-frame-server*, and the *sticky-note-displayer*. We describe the first two components below, and the final is covered in Sect. 4.
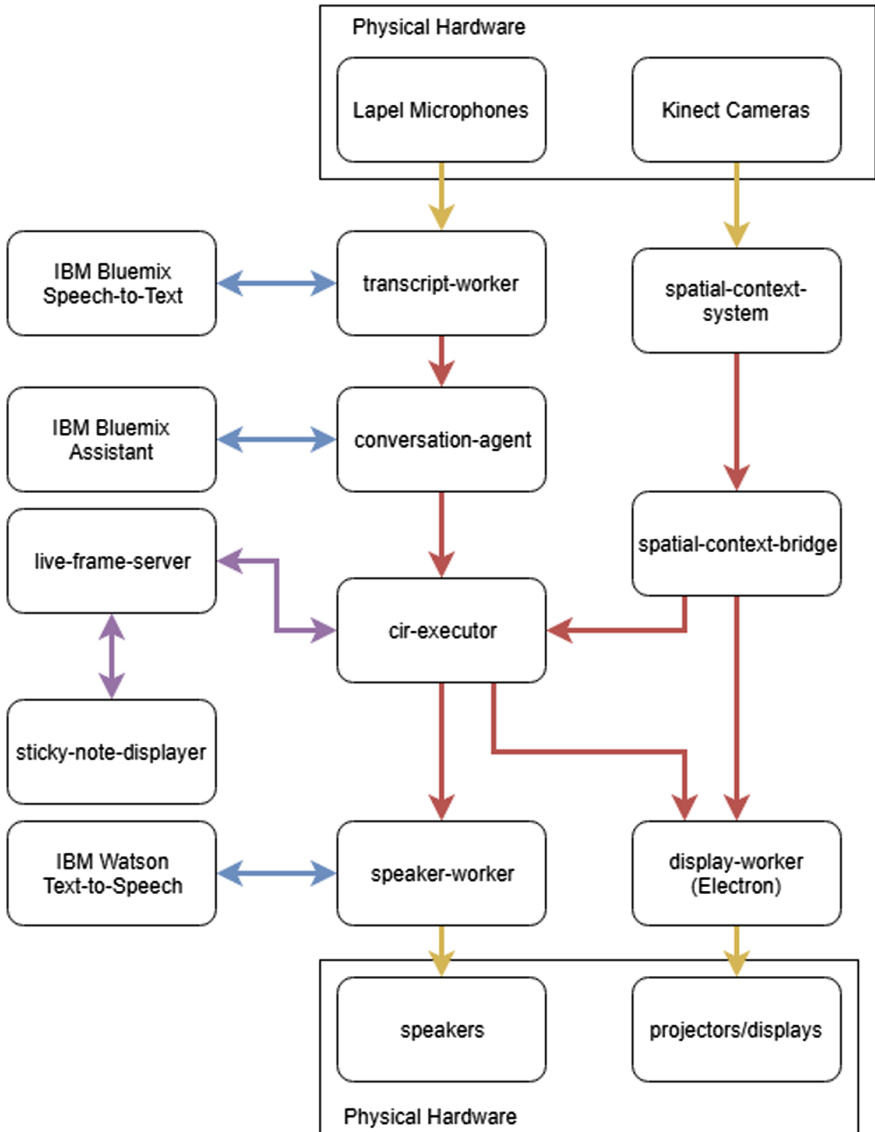
Fig. 1. Diagram of the cognitive and immersive system

3.1 The Spatial-Context-Bridge

The *spatial-context-bridge* takes input from the spatial context system [11] that creates a semantic encoding of the data the kinect sensors are receiving. This positional and gesture data is then used to alert the executor when a gesture changes (e.g. going from an open hand to a closed hand) as well as to push the relevant information to the display-worker to show the user's hand positions

within Electron. This is all done without wrapping ourselves the physical mouse interface of the computer. While prior work showed the power of doing that [3], it created a limitation of a single user at a time of the system. Instead, we utilize a concept of a "virtual mouse" that exists per user, which does not use the actual mouse interface of the computer platform. To do this, we make use of the rich scripting API that the web affords us through JavaScript and Electron framework. In this case, we can utilize the MouseEvent APIs to simulate mouse clicks, scrolling with the mouse, holding the mouse button down and releasing it, etc. Within a CAIS, if one, while pointing at the screen, closes their hand, it triggers a mouse down event in the system. If that person was to open their hand, it triggers a mouse up event in the system, and if the down/up was quick enough together (and in similar position), triggers a click event. While our system does support other actions, we omit them here as they are not used for our tool. For each of these generated actions, we build in EventListeners within the page to react in some way when one of these actions are detected (whether it's done through the physical mouse, or through our virtual mouse) and to take an appropriate action. In the case of the virtual mouse events, as part of the payload describing the event, we attach an ID of the person performing the event, so that we could then link chains of mouse events to the same person, and thus allow multiple users making discrete actions without interfering with each other. Unfortunately, this does require any webpage that we want to use as multiple users to be specially programmed to take advantage of our extra metadata in the events. However, for any webpage that does not, we are equal to prior work in that as we're mimicking the events a real mouse does, the pages are still fully usable by a single person through the same kinect based interface as a mouse. Additionally, our specially designed webpages are also usable by a single person on a regular mouse as well by that same token.

3.2   The Live-Frame-Server

The *live-frame-server* gives us a distributed data pattern such that it enables many clients to connect to a central server which keeps them in-sync. The server holds many independent "frame", each with their own unique ID and are backed by MongoDB to allow for picking back up a frame in later sessions within the CAIS. Each frame is composed of a JSON object that any connected client can update. Each client connects to the server via websocket passing in a specific frame ID that it is interested in. When one client sends an update to the server, the server calculates the diff of the change, and then pushes both that diff as well as the new JSON object for the frame to all other clients, keeping everyone in-sync with all changes. The final component is the *sticky-note-displayer* which houses the display logic for our clients. Each client that connects through the displayer is connected to the same frame, allowing us to give multiple views as well as allow users to utilize their own personal devices.

4   The Digital Brainstorming Tool

The digital version of the brainstorming tool was designed to try to follow the traditional pen-and-paper process as closely as possible, while taking advantage of the additional capabilities that the CAIS provides. To accomplish this, we drew mock-ups and designs for how the high level actions taken in the pen and paper process in a digital tool that took advantage of the space offered by the CAIS. To accomplish this, we developed two separate interfaces for interacting with the system. The first is the "personal view" which is displayed to users via a personal device (such as laptop or tablet), shown in Fig. 2 and is linked to a specific user (though the same user could connect to the "personal view" on multiple devices at the same time).

Within the view, the user sees all categories that have been created within the system as well as all of the notes belonging to each category. The first section however contains a "personal" category such that any notes within it are only viewable and editable by that user. Users are free to create as many categories as they want, as well as many notes as they want per category. The users can then delete any category, which would include the notes in that category, except for their personal category. Any created category or note within any of the non-personal categories are then synced to any other connected "personal views". The interface is used through tapping on the icons (such as the plus icon within a category to create a note, or the x icon to delete the category) and the keyboard. When a user clicks to create a category or note, a prompt appears which allows the user to type into it to fill the category title or note content. To move a note from a category, the user clicks on the note, and then clicks on a category to move it to.

The other interface for the system is the "global view", shown in Fig. 3. This view was split in two with the left half used for the notes in the "uncategorized" category, and the right half containing all the of the created categories and the notes within those categories. Interaction with this display is done through a combination of voice and gestures for commands, as well as picking up any changes done via any open "personal view". The allowed voice commands are:

1. Create a note that reads "text"
2. Create a category named "title"
3. Rename the category named "title" to "new title"
4. Move note "ID" to category named "title"
5. Delete note "ID"
6. Delete category named "id"

For commands 3–6, the one of the blanks can be substituted by a pointing gesture at the item you're attempting to operate on. An example of would be saying, "Move that note to category named foo." In parsing this command, the executor is given a move note intent by the user which requires a note and category, however the speech only supplies the category. The system then contacts the "spatial-context-bridge" to find out where that user is currently pointing, which is given as an x-y coordinate on the screen. The executor then
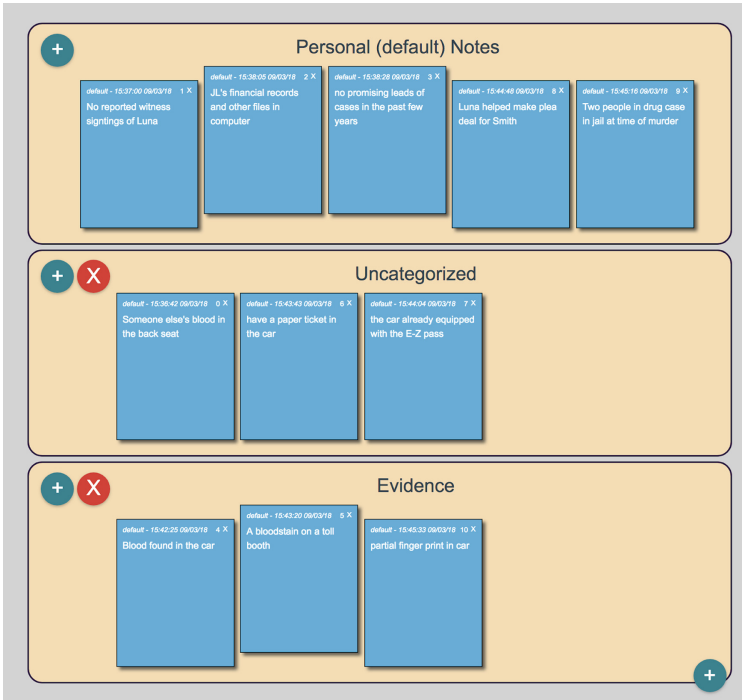
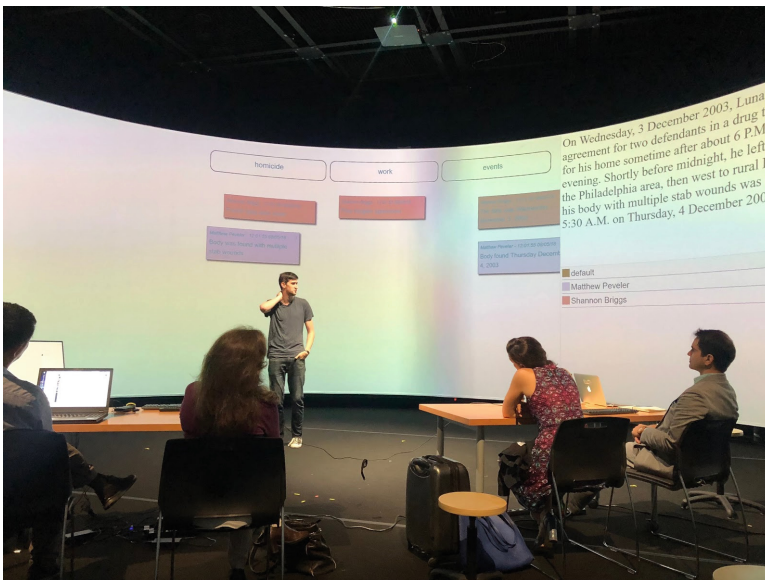**Fig. 2.** Personal view of brainstorming tool

Fig. 3. Global view on panoromic display

translates that into a relative x-y coordinate within a given webview, which is then used to get the elements that are at that coordinate within the webview, and to check if any of the elements match the type we're looking for. By getting all elements at the coordinate, and not just the top-most, we can handle resolution of moving a note via its ID while pointing at an existing note within a category, such that the category element is underneath the note element we're pointing at, but we're referencing the category.

5   User Study

An ongoing user study uses participants from a local university to test the usability of our digital brainstorming tool in comparison to the more traditional pen and paper tool. The study is run via a within-subjects design where participants go through two brainstorming sessions in groups of three. In both sessions, the participants go through the same brainstorming process, which is broken up into several five minute sections. To start with, the participants are given the same one-page excerpt of a murder case to read. Following the reading phase, participants are asked to individually create notes about the most relevant details of the excerpt. The next phase involves the participants coming together as a group and then putting all of their created notes into a shared workspace. After all notes are in the workspace, the participants are asked to start grouping the notes based on similarity into categories. Finally, the participants have a discussion phase about the case, as well as if they feel that they may have missed any important details in the grouping, or if there are some notes that do not fall into one of their categories. To finish the session, the participants are then asked to fill out a brief survey as well as conduct a question and answer session on what they thought about a given tool. The first session involves using the pen and paper approach as described in Sect. 2. The second session involves using the digital tool, both the overall CAIS and global workspace, as well as the personal view on a tablet that each participant has as described in Sect. 4.

6   Discussion

The preliminary user studies conducted gives us valuable feedback on how we may improve the tool and its functionality. We conducted two rounds of our study, implemented improvements to the system based on their feedback, and then conducted another round.

From the first two rounds, the biggest criticism that we received was that the tool was not as easy to grasp as we had originally supposed going in. While we provided users with a brief introduction to all technology, many felt it hard to remember all of the commands to the system, as well as how to mix voice and gestures. The feedback from the users was that they would have liked an extended video, tutorial, or sheet of commands to the system. Following this, we implemented a view to the right of their primary workspace that listed all commands that could be issued to the system with some slight variations in how

they could say it. The final group found this useful and was able to use the system more fluently without having to ask us questions on commands.

When using the tool, while participants did find the live transcription of what they were saying useful, they were consistently frustrated by the strictness of some commands as they attempted to leave out words from the phrases we had inputted. An example of this is the command "Move the note into the category named foo" to move a note. Users would often times leave out the "named" part of the command, which messed up our parser in finding the category name in the sentence. However, the system did not make it obvious that the command was not properly parsed, and users were left wondering what had went wrong. In addition, users identified various features that they thought were missing from the digital tool. The most significant feature was the ability to undo a deletion of a note or category. Whereas in the pen-and-paper version, removed notes were just placed off to the side, and could be quickly and easily re-put up on the wall if necessary, notes that were removed from the digital tool had to be recreated from scratch. Categories that were deleted that contained notes would also delete all notes underneath that category, which surprised users who were expecting the notes to automatically go back to being "uncategorized".

With the gesture system, we found mixed results with participants. Several found the system easy to use, while others found it difficult. The biggest issue we found was that the system was designed for a person to be gesturing with one hand while their other hand was at their side. When that other hand was holding a tablet, it caused the system to mistakenly classify the hand as holding the tablet as a gesture in addition to the hand actually making a gesture.

Most startling was the mixed feedback on the collaborative benefits of the digital tool with some subjects stating that they felt more involved in the discussion and another stating that they felt disconnected from the other participants. It was not clear with the negative response whether their disconnect from other participants was principally due to the platform, or mostly do the problems they had with using the tool. For future user studies with an improved version of the tool, it will be important to monitor engagement of participants in the group discussion phase of the digital tool session compared to the pen and paper process, and attempting to pinpoint why a disconnect might exist.

7   Conclusion

We now quickly summarize our work and present promising lines of future work. Our primary contribution is in the development of an integrated sticky notes application used for brainstorming which easily allows for both individual and group usage. Our paper describes contributions that we've made in components necessary in a cognitive and immersive system to support this work, as well as for future work, especially for the design of the virtual mouse interface. This provides a generalized interface that can be attached to any use-case within our cognitive and immersive system in both providing for interaction via gestures of webpages, both specialized to take advantage of normal users, as well as normal webpages supporting single users.

However, as shown in Sect. 6, thoughts of our tool by users were mixed. At this stage, we plan to look to address many of these concerns, before conducting future user studies as we continue to believe that the digital tool has the potential to become invaluable to this process. To accomplish this, we plan to add in a new "trash" section within the tool that stores notes that have been deleted, and through which users can recover notes. For categories that are deleted, any note under it should go into the "uncategorized" section of the screen, instead of being deleted as well. A more substantial change is in loosening the parsing of commands to allow for more formats of each command (as well as missing words in the command), while also adding in the capacity of our system to prompt the user for information it could either not find in the command or did not understand. An example of this is that the system would sometimes incorrectly transcribe someone when they attempted to say a category's name, and then do nothing, which would force the user to look at the live transcript and parse it to see what went wrong. The system should at the very least prompt the user for the information it was missing, giving the user a notion that their previous command got messed up, as well as just requiring the user to say the missing information, and not the entire, potentially lengthy command. In addition, for commands that referenced a category, if the category name did not exist, the system could then check to see if any of the existing names were similar (an example of this would be using the levenshtein distance) and if so, prompt the user to see if that was what they meant, requiring just a yes or no response from the user. This would hopefully help to relieve frustration with the system when it messed up, as well as make it clearer to the user what went wrong. In addition, we plan to look into how we may allow users to better utilize the digital format in attaching non-text content to their notes as part of the brainstorming process, and how this content could then be showed to the users on demand. A primary consideration as well is needed in making the tool feel more intuitive to the users such that using it does not detract from the sense of discussion of the topic at hand.

Future work will consist of expanding the work to be used to drive future stages of a sense-making model [8]. The next stage of the model is that of dealing with alternative competing hypothesises, utilizing the notes created in the brainstorming session for evidence that might support or contradict any created hypothesis. In addition to the sense-making direction, we hope to apply the sticky notes application as something we can utilize within other domains to help capture whatever it is that users are discussing and to help them categorize and share their thoughts with their colleagues.

References

1. Baber, C., Attfield, S., Conway, G., Rooney, C., Kodagoda, N.: Collaborative sense-making during simulated intelligence analysis exercises. Int. J. Hum.-Comput. Stud. **86**, 94–108 (2016)
2. Beebe, S.M., Pherson, R.H.: Cases in Intelligence Analysis: Structured Analytic Techniques in Action. CQ Press, Thousand Oaks (2012)

3. Brooks, R.A.: The intelligent room project. In: Proceedings of the 2nd International Conference on Cognitive Technology (CT 1997), p. 271. IEEE Computer Society, Washington (1997)

4. Divekar, R.R., et al.: CIRA: an architecture for building configurable immersive smart-rooms. In: Arai, K., Kapoor, S., Bhatia, R. (eds.) IntelliSys 2018. AISC, vol. 869, pp. 76–95. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-01057-7_7

5. Dornburg, C.C., Stevens, S.M., Hendrickson, S.M.L., Davidson, G.S.: Improving extreme-scale problem solving: assessing electronic brainstorming effectiveness in an industrial setting. Hum. Factors **51**(4), 519–527 (2009). pMID: 19899361

6. Jensen, M., Thiel, S., Hoggan, E., Bødker, S.: Physical versus digital sticky notes in collaborative ideation. Comput. Support. Coop. Work **27**(3–6), 609–645 (2018)

7. Mistry, P., Maes, P.: Intelligent sticky notes that can be searched, located and can send reminders and messages. In: Proceedings of the 13th International Conference on Intelligent User Interfaces IUI 2008, pp. 425–426. ACM, New York (2008)

8. Pirolli, P., Card, S.: The sense making process and leverage points for analyst technology as identified through cognitive task analysis. In: Proceedings of International Conference on Intelligence Analysis (2005)

9. Saleh, M., Abel, M.H., Misséri, V., Moulin, C., Versailles, D.: Integration of brainstorming platform in a system of information systems. In: 8th International ACM Conference on Management of Digital EcoSystems (MEDES2016), pp. 166–173. Hendaye, France, November 2016

10. Widjaja, W., Yoshii, K., Haga, K., Takahashi, M.: Discusys: multiple user real-time digital sticky-note affinity-diagram brainstorming system. Procedia Comput. Sci. **22**, 113–122 (2013)

11. Zhao, R., Wang, K., Divekar, R., Rouhani, R., Su, H., Ji, Q.: An immersive system with multi-modal human-computer interaction. In: The IEEE Conference on Automatic Face and Gesture Recognition, May 2018