An Experiment in

Machine-Augmented Sensemaking

in Intelligence Analysis

ABSTRACT

Singapore has developed a prototype Risk Assessment and Horizon Scanning (RAHS) system in collaboration with The Arlington Institute and Cognitive Edge, with the aim of providing analysts with an extensible suite of tools for data exploration and data exploitation based on a service-oriented architecture. The RAHS system facilitates the extraction of open source information into repositories, which are then made available to analysts for search and retrieval by means of various tools to augment the users' sensemaking process. This paper describes how the RAHS system may be used in the analysis of a massive amount of data. In addition, this paper also presents a limited experiment in which analysts were presented with the task of exploring a set of documents related to a fictitious terrorist organisation in order to identify the roles and responsibilities of the various people linked to the terrorist organisation. A simple comparison regarding the workflow, efficiency, and effectiveness of a RAHS analyst is made with that of an analyst equipped with a traditional search engine. Finally, the lessons drawn from this experiment point to the need to cultivate a team of analysts who would be able to leverage the flexibility of the RAHS system in supporting a balanced heuristic.

Gwenda Fong
Keith Oh

An Experiment in Machine-Augmented Sensemaking in Intelligence Analysis

INTRODUCTION

For most agencies in the business of surveillance or other forms of research, open sources of intelligence present a promising and vast ocean of information for the analysts to trawl through because of the massive amounts of relevant data-points the search could potentially amass in a relatively short period of time. For instance, eyewitnesses posted accounts of the London terrorist bombings in July 2005 on Wikipedia and blogs just minutes after the incident (Thompson, 2006). Further updates then flowed in over the next couple of hours ("Timeline of the 2005 London bombings", 2007). The police culled ground intelligence inputs from closed-circuit TVs installed in public areas and collected pictures and movies sent in from eyewitnesses' mobile phones ("Report of the official account of the bombings in London on 7th July 2005", 2006; "Police appeal for assistance - can you help", 2005). The richness of data from multiple channels enabled the police to shortlist and apprehend the culprits swiftly.

In other domains such as scientific research where time is not as critical, rich databases containing conference proceedings, journal articles, online books and the like are also readily available at the click of a button, although sometimes at a premium. Even the intelligence community is looking at using Intellipedia (Thompson, 2006) to share information and their assessment of the information. Of course, members of the intelligence community will first have to overcome the dissonance between the practice of information sharing and the covert nature of their operations, as well as consider other legitimate concerns regarding the security of sensitive data and social ethics that often plague intelligence work, before Intellipedia can be operational. Clearly, the common challenge across the different domains of surveillance and research goes beyond data collection; the challenge lies in enhancing the analyst's subsequent sensemaking process as he finds himself faced with a glut of information.

Theoretical Models and Previous Studies

A number of researchers have studied the process of intelligence analysis in terms of

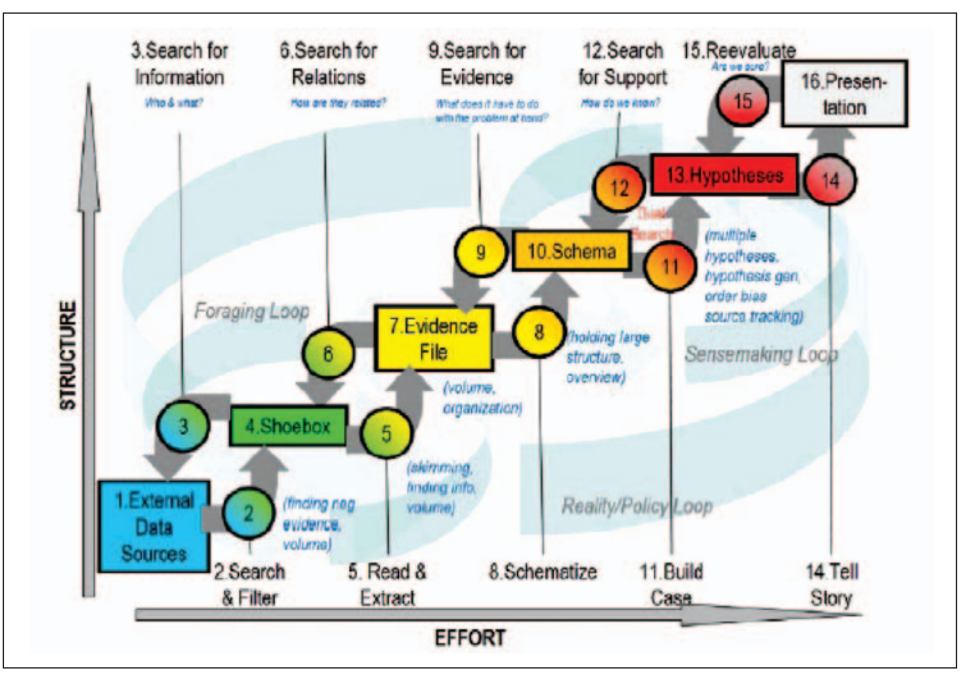

Figure 1. Pirolli and Card's model of sensemaking for intelligence analysis

sensemaking. While it is not the aim of this paper to present detailed descriptions and comparisons of the various theoretical models and studies done by these researchers, it will highlight some theories that pave the way for machine-augmented sensemaking.

Pirolli and Card (2005) have broken down the analyst's foraging process by using cognitive task analysis methods (See Figure 1). The subtasks are a series of information foraging and sensemaking loops that are necessarily iterative and tightly coupled to reflect the complexity of the analyst's thought process in making sense of the data.

Klein et al. (2004; 2006a; 2006b) proposed an alternative Data/Frame model of the nature of sensemaking activities. The model emphasises the creation and iteration of frames, which represent the analyst's perspectives in his or her deliberate effort to understand events (See Figure 2).

In both approaches to explaining the analyst's process, the analyst goes through both a top-down process where predetermined hypotheses drive information searching, as well as a bottom-up process where information searching drives hypotheses formation (Bodnar, 2005). Hypotheses are thus constantly proven or refuted during sensemaking.

The process of verifying hypotheses is an automatic process for expert analysts. Experts operate with a high level of automaticity because their skills have been converted into procedural knowledge (Anderson, 1982). Ericsson et al. (1993) investigated the role of "dedicated practice" in acquiring expertise; the belief is that it takes a minimum of 10 years to become an expert. Must it always take at least 10 years of "dedicated practice"? Ericsson and Lehmann (1996) found schemas developed in experts, not unlike the procedural knowledge that explains automaticity in experts. However, these schemas have to be more than random patterns of the elements in the task. Chase and Simon (1973) studied chess experts and discovered that chess experts could recall visual patterns of meaningful chessboard configurations but not random configurations. In a digit recall task (Chase & Ericsson, 1982), SF, though an ordinary college student, was able to increase his digit span for recall from seven to 80 after 230 hours of practice. Chase and Ericsson attributed SF's feat to his strategy in relating those numbers to running timings. It should be no surprise that SF was an avid long-distance runner. SF has shown that it is possible to short-circuit the pre-requisite of 10 years of "deliberated practice" for a memory task given the appropriate strategy. A powerful memory aids analysis work, as the analyst will be able to

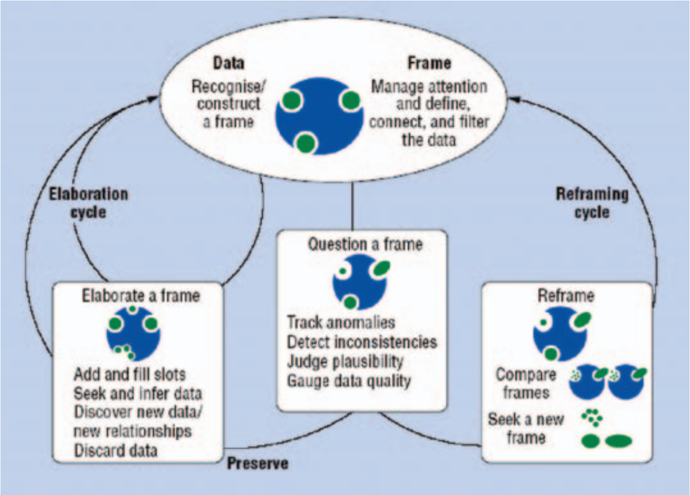

Figure 2. Klein's Data/Frame model of sensemaking

recall data readily to verify hypotheses. The key strategy is to focus on the relationships among the data, rather than the data itself. Meaningful relationships among data are parallel to the structures of long-term memory and will thus facilitate chunking, the learning mechanism through which information is acquired (Chase & Simon, 1973). The consistent theme across the sensemaking literature calls for a system that presents not only data, but also metadata and meaningful relationships among the data and metadata to the analyst.

With such a direction in mind, the Risk Assessment and Horizon Scanning (RAHS) system was designed to support non-prescriptive and flexible workflows where meaningful patterns in the data and metadata are presented to the analyst in various forms of visualisation. Interaction with the visualisations would enable the analyst to further refine the patterns. The aim is not to negate the need for competent analysts, but to free up the analysts' time from manipulating the data, so that time can be better spent on analysing the data and the patterns within the data.

RISK ASSESSMENT AND HORIZON SCANNING

DSTA has developed a prototype RAHS system in collaboration with The Arlington Institute and Cognitive Edge. Features available in RAHS include:

- **Search** - Apart from a basic keyword search, the search engine allows the analyst to expand the search terms based on semantics and patterns in spelling. This helps to cast a wider net for relevant documents and avoid a situation where documents are missed out due to different terminologies and spelling variations. In addition, the Boolean search syntax allows the analysts to specify some logic rules among the keywords (e.g. AND, OR, NOT, etc.) to further define a broader or narrower search criterion. Providing an effective search is important; we speculate that it is one of the most common entry points into the system and it determines the relevance of the data entering the "shoebox" in the foraging loop (Pirolli & Card, 2005).

- **Clustering** - Themes are extracted from documents as metadata. A set of documents (e.g. search results, folder contents) can be grouped according to the similarities in theme. This provides a thematic approach for the analyst to organise the data in his shoebox; it is especially useful for the analyst to isolate his attention to a theme specific to his research interests for the moment.

- **Duplicate Detection** - Due to the nature of media, various news publishers can publish the same news article. A search will inevitably return documents from different publishers, but with the same content. The shoebox should contain only data that is essential for evidence extraction; otherwise, the analyst will be bogged down with redundant data (unless the fact that various publishers carry the same news article is also of interest). This is a systems issue that can be eradicated by Duplicate Detection when it compares the contents of the documents to determine if they are duplicates.

- **Summarisation** - Summarisation is needed in many processes. In the later part of the foraging loop where the analyst has to extract evidence from the shoebox, the analyst has to read a lot of new data and constantly review the contents in his shoebox. Summarisation picks out key sentences from documents; this allows the analyst to get the gist of the new data, and provides a quick preview to what is already in the analyst's shoebox.

- **Keyword Analysis** - The analyst may have a set of keywords that he monitors. Keyword analysis provides a visualisation of how these keywords are mentioned in the data. When the analyst specifies the set of keywords for monitoring, he is implicitly building a schema around the elements in the problem. He believes that these elements have a role to play in the problem, and that the mention of these elements in the data should behave in a certain manner. Conformity to the hypothetical behaviour reinforces the schema, while deviation from it challenges the analyst to introspect further on his presumptions.

• **Entity Analysis** - Using natural language processing, the system extracts entity metadata, e.g. people, organisations, locations from the data, and provides a visual representation of the occurrences of these entities. Further plots of how these occurrences vary over time and how a single entity is associated with the rest are also available. An interesting point to note is that this was essentially how Dan Swanson manually discovered that magnesium deficiency contributed to migraine (Swanson, 1987). Swanson basically conducted a cross-sectional analysis of entities associated with migraine across the literature and identified "spreading depression" as one entity with a higher occurrence. He repeated the process to get the second order of entities associated, and came to "magnesium deficiency". This provided a starting point for him to schematise how magnesium deficiency and migraine can be possibly related.

• **Entity Network** - Taking entity analysis a step further into the information foraging loop, the entity network draws relationships among the entities. This is presented in a network diagram of nodes and links which the analyst can export to Analyst Notebook (Analyst notebook product overview, 2007) for editing.

• **Timeline Analysis** - Events are extracted from the data, based on user-specified keywords and are arranged in a chronological order. This is a familiar presentation of events not unlike what we see in history textbooks. In some cases, relationships among the data are only apparent when presented in a temporal manner, such as cyclical occurrences of events. The timeline is thus a possible channel to obtain evidence for such a schema.

• **Automatic Questioning and Answering -** During any part of information foraging or sensemaking, the analyst might need a quick question answered, e.g. "Who is John G. Doe?" The system provides a fact-based question-and-answer (Q&A) module where it searches the database of documents for relevant sentences that best answer the question. The functionality is also part of entity extraction in that the analyst can ask several "who is" questions about the people entities in one go. This helps to short circuit the path from the sensemaking loop back to the information foraging loop.

THE EXPERIMENT

A system and concept demonstration of a pre-beta version of RAHS was held in conjunction with a command post experiment that was jointly conducted by the Singapore Armed Forces (SAF) and the US Defense Advanced Research Projects Agency (DARPA) in October 2006. A dry run was held in August 2006, during which a traditional search engine capability was fielded to provide a baseline comparison with the intelligence analysis workflow facilitated by RAHS. The search engine used during the dry run was a commercial enterprise search solution employed by a variety of companies and industries. It provided users with the generic basic and advanced search functions, and would return the list of documents that satisfied the search criteria. The remainder of this paper describes the experimental methodology, as well as the results and a brief discussion of this Limited Objective Experiment.

Participants

The system and concept demonstration of RAHS was designed as part of a larger experiment involving approximately 130 SAF and US Army officers (spread over two runs) who were tasked to role-play a coalition joint task force in the October experiment. The RAHS intelligence analysts were featured as a small component of the coalition task force, whose roles were played by two SAF officers who had just completed a course with the Singapore Command and Staff College,

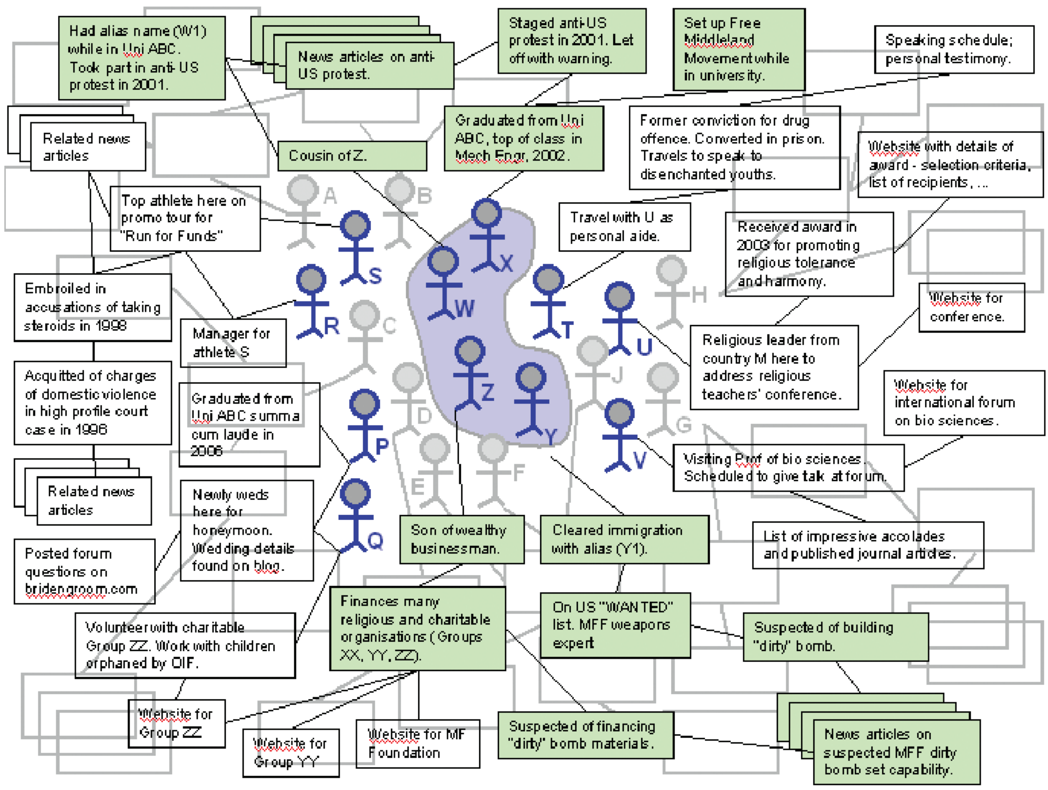

Figure 3. Graphical depiction of repository of documents developed for the experiment

preparing them for command positions. In contrast, the role of intelligence analysts using only a traditional search engine in the August dry run was played by two junior SAF officers from the SAF Centre for Military Experimentation (SCME).

We believe that the difference in seniority between the two groups of participants is mitigated by the fact that the task assigned to them did not require them to draw on any domain-specific expertise or experience. In fact, the RAHS analysts were likely to be slightly handicapped in that they had to learn to use a new tool as well as a new search paradigm, whereas the analysts using the search engine had the advantage of being familiar with the tool and the traditional search heuristic. However, sufficient training was provided to the RAHS analysts to familiarise them with the system so as to eliminate biases due to unfamiliarity.

Scenario

The scenario developed for the purpose of this experiment centres on UN-sanctioned coalition operations in the fictitious countries of Jurongland and neighbouring Middleland. The scenario starts with a picture of the Middleland military's occupation of an unoccupied former US Naval facility in Jurongland. The UN has asked Singapore and US forces to form a coalition joint task force to establish a Zone of Security to prevent further invasion by the Middleland forces, and eventually restore territorial integrity to Jurongland. In addition to the traditional military threat posed by the Middleland forces, the coalition task force also has to contend with a group of militia who call themselves the Middleland Freedom Fighters, and who have aligned themselves to the Middleland government's objective of reunification with Jurongland to form a Greater Middleland.

Task

While the coalition task force was occupied with the traditional military planning process, the analysts were tasked to search through a database of over 100 documents, approximately 10% of which contribute information to the piecing together of the ground truth of who was affiliated to the Middleland Freedom Fighters, their relationship to one another, as well as their roles and responsibilities. However, the majority (90%) of documents in the repository actually serve to decrease the signal-to-noise ratio. These documents were related in some way to the main topic of interest (e.g. in terms of related key words, as well as people who are involved with the same organisations but not in a criminal way), so as to confound the matter. Figure 3 shows a graphical depiction of the repository of documents developed for the experiment. It shows that the relevant documents (highlighted in green), taken together, would point to the involvement of persons W, X, Y, Z in the fictitious terrorist organisation.

Measures

The success indicators used to assess the utility of the RAHS system and concept as compared to that of a traditional search engine were identified to be: (a) accuracy of intelligence analysis, and; (b) workflow process. These were measured using a combination of participant questionnaires as well as periodic screenshots. A description of how each success indicator was measured is given here:

- **Accuracy of intelligence analysis** – The officers who took on the role of intelligence analysts were asked to answer a series of 10 questions relating to the scenario that they were presented with, approximately 150 minutes after the start of the experiment run. They were not allowed to use RAHS or the search engine to conduct further research while answering these questions. These questions were designed to assess how far each analyst had advanced in his analysis of the glut of information that he had been presented with. The questions posed to the analysts comprised six "strong signal" questions (i.e. questions to which answers could be found by reading a single document) as well as four "weak signal" questions (i.e. questions to which answers could only be inferred from reading two or more documents). An example of a "strong signal" question was, "Who is the Middleland Freedom Fighters' weapons expert?" The answer to this question could be easily found by reading a description of a hypothetical FBI's list of wanted people. In contrast, an example of a "weak signal" question was, "Which members of the Middleland Freedom Fighters are related and how are they related?" The answer to this question could only be found by linking information present in three separate documents – firstly, one that establishes the link between person A, the son of a well-heeled entrepreneur, with the Middleland Freedom Fighters, secondly, one that establishes the link between person B and the Middleland Freedom Fighters, and lastly, a third seemingly innocuous document that identifies person B as the nephew of the same well-heeled entrepreneur, thus making persons A and B cousins. The questionnaire was deliberately designed to include both "strong signal" as well as "weak signal" questions in order to assess if the RAHS system would better facilitate the analysts in detecting weak signals in addition to flagging out the obvious "strong signals".

- **Workflow process** – The workflow processes of both sets of analysts were captured through a series of screenshots taken periodically at between three-second and six-second intervals. The post-hoc analyses of the screenshots elucidated descriptive statistics of the various usage patterns such as the frequency as well as average duration each analyst spent on each feature of the respective tools that were made available to them.

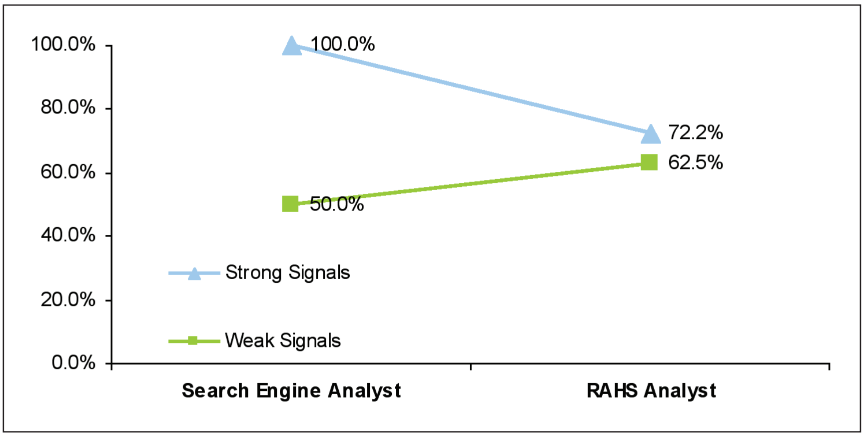

Figure 4. Accuracy of intelligence analysis as facilitated by the Search Engine vs RAHS

RESULTS AND DISCUSSION OF FINDINGS

Accuracy of Intelligence Analysis

Figure 4 shows the accuracy of intelligence analysis as facilitated by a traditional search engine capability when compared with RAHS demonstrated in the experiment.

Clearly, the Search Engine analyst outperformed the RAHS analyst in terms of the percentage of strong signals detected, while the RAHS analyst outperformed the Search Engine analyst in terms of the percentage of weak signals detected in this simple experiment.

A closer look at the two workflows explains why this was so. The search strategy employed by the Search Engine analyst may be thought of as a narrow but in-depth paradigm. It is narrow because the analyst has to read the individual documents returned as a result of his search query; consequently, there is only a limited number of documents that the analyst can cover under time pressure, and any cross-document inference of information is largely dependent on the analyst's understanding and memory of what he has read. However, this strategy affords the analyst the benefit of appreciating the context of each document that he reads. This is reflected in the descriptive statistics of the Search Engine analyst's accuracy of intelligence analysis – while the analyst was able to answer correctly all the "strong signal" questions (i.e. questions to which answers could be found by reading a single document), the analyst did poorly in answering the "weak signal" questions (i.e. questions to which answers could only be inferred from reading two or more documents).

In contrast, the search strategy employed by the RAHS analyst may be described as a broad but cursory coverage of the database of documents. It is broad because the automated question and answer feature allows the analyst to pick out only the key sentences across the subset of documents related to the search query, thus facilitating the analyst's linking of information across several documents. However, the drawback is that the analyst may become overly reliant on the summaries provided for him and neglect to read the source documents from which the sentences are picked out, thereby missing out on the larger context in which the statements are made. This provides an explanation for the descriptive statistics of the RAHS analyst's accuracy of intelligence analysis in this experiment – the analyst missed out on answering some of the "strong signal" questions as compared to the Search Engine analyst, but did better in terms of answering the "weak signal" questions.

It should be stated as a caveat that it is less important to focus on the absolute percentages as a description of the analysts' performance (i.e. that the Search Engine analyst detected 100% of the "strong" signals) than to compare their performances across the two sensemaking systems. This is especially so because of the small scale at which the experiment was conducted (with a database of slightly over 100 documents), as compared to real life open source information on the Internet, for instance. It is expected that the percentages of "strong" and "weak" signals detected would be somewhat lower with a much larger repository of information to search from, especially if the analysts only had access to traditional search engine capabilities.

The lesson to learn from this simple experiment is that there is a need to balance both these intelligence analysis strategies. The experiment has shown that there are benefits and drawbacks associated with each search paradigm, and that it would be prudent to develop a heuristic that harnesses the advantages of both paradigms. At this point, it should also be mentioned that the RAHS system comprises a powerful search engine capability as part of its suite of tools, but the RAHS analysts in this experiment were deliberately steered away from using the traditional search paradigm with the main objective of seeing large differences in outcome between the two groups of users. However, as the RAHS system and concept of use develop and mature, a team of RAHS analysts will be well positioned to develop a heuristic that leverages the strengths and benefits of both paradigms to facilitate their sensemaking of massive amounts of data.

Workflow Process

Tables 1 and 2 show the descriptive statistics regarding the various usage patterns as engendered by a traditional search engine and that of the RAHS system.

Table 1 shows the descriptive statistics on the Search Engine analyst's workflow. Discounting the miscellaneous activities the analyst was engaged in (e.g. infohub, map, chat, email, reading of breaking news), the statistics reveal that the analyst spent the most time (42.3 minutes) reading the repository documents returned by the search function. A significant amount of time was also spent performing the search itself (22.1 minutes), as well as using an excel spreadsheet to compile the relevant information gleaned from reading the repository documents (15.8 minutes). The statistics on the frequency of screenshots for each category as well as the average time spent on each activity indicate that the analyst was constantly switching between activities, spending only an average time of between 10.6 and 15.0 seconds on each activity at a stretch.

Screenshot Activity	Frequency in approx. 150 min	Average time spent	Total time spent
Search	125	10.6 sec	22.1 min
Reading of repository documents	169	15.0 sec	42.3 min
Excel spreadsheet compilation	75	12.6 sec	15.8 min
Reading of breaking news	53	13.5 sec	11.9 min
Others (infohub, map, chat, email)	169	19.5 sec	54.9 min
Screensaver	0	–	–

Table 1. Descriptive statistics on the Search Engine Analyst's workflow

Screenshot Activity	Frequency in approx. 150 min	Average time spent	Total time spent
Login	2	39.0 sec	1.3 min
Main Menu	3	21.0 sec	1.1 min
Search	6	157.0 sec (2.6 min)	15.7 min
Reading of repository documents	0	–	–
Timeline analysis	1	42.0 sec	0.7 min
Entity analysis	7	164.7 sec (2.7 min)	19.2 min
Automatic Q & A	18	71.7 sec (1.2 min)	21.5 min
Word document compilation	19	140.5 sec (2.3 min)	44.5 min
Others (news, email, etc)	6	121.0 sec (2.0 min)	12.1 min
Screensaver	2	423.0 sec (7.1 min)	14.1 min

Table 2. Descriptive statistics on the RAHS Analyst's workflow

Table 2 shows the descriptive statistics on the RAHS analyst's workflow. Discounting the miscellaneous activities that the analyst was engaged in (e.g. reading news, email), the statistics reveal that the main activities that preoccupied the analyst were: compiling relevant information in a word document (44.5 minutes), automatic Q&A (21.5 minutes), entity analysis (19.2 minutes), and search (15.7 minutes). It is noteworthy that the analyst did not choose to read any of the repository documents in their entirety even though they were available. Instead, he had chosen to rely completely on the summaries provided by the automatic Q&A. This, again, lends support to the description of the RAHS analyst's search strategy as one that is broad but cursory, and thus vulnerable to missing out "strong" signals as described. The statistics on the frequency of each category of screenshots, as well as the average time spent on each activity, paint a rather different picture from that of the Search Engine analyst's workflow. While the Search Engine analyst was constantly switching between activities and spent only a short amount of time on any activity at a stretch, the RAHS analyst, on the other hand, switched between activities a lot less frequently, and tended on average to spend a much longer amount of time (between 1.2 and 2.7 minutes) on each activity at a stretch.

The fact that the RAHS analyst was not occupied with a frenzied switching between activities, unlike the Search Engine analyst, may be taken as a good sign that the RAHS system had automated and collapsed some of these iterative information foraging loops as displayed by the Search Engine analyst. This was largely attributed to the RAHS system's ability to facilitate entity analysis across the subset of documents returned from a search query, and the automatic Q&A feature that made salient the entities of interest. However, the flip-side was that there was a tendency for RAHS analysts to focus on a single workflow. Although it provided some initial answers early in the intelligence analysis process, it soon reached a saturation point and did not return any new pieces of the puzzle. The RAHS analysts should ideally have challenged themselves to innovate on their workflow constantly and expand their information foraging strategy to gain new insights to the problem, instead of getting lulled into a sense of complacency and becoming overly reliant on a single workflow to recover all the pieces of the puzzle.

CONCLUSION

The purpose of this paper was to describe an approach to machine-augmented sensemaking applied specifically to the area of intelligence analysis. This was explored in a Limited Objective Experiment jointly conducted by SCME and DARPA in October 2006. The purpose of this experiment was to demonstrate the RAHS system and its concept of use, as applied to intelligence analysis in the context of a command post experiment. A traditional search engine capability was also fielded in a dry run in August 2006 to provide a baseline comparison to the intelligence analysis workflow facilitated by RAHS.

The small number of participants in this experiment does not allow us to make meaningful statistical comparisons of their workflow and performance across the various sensemaking tools. Nonetheless, the data collected elucidated significant differences in the workflow processes as facilitated by the respective tools, which would likely have contributed to the differences in the quality of intelligence analysis with regard to "strong" and "weak" signals.

The lesson that we have drawn from this system and concept demonstration is that the desired outcome for the SAF is to have a team of analysts develop a balanced heuristic for more effective sensemaking of massive amounts of data. It would also be prudent to guard against over reliance on a single workflow that may lull the analysts into a sense of complacency and stifle any innovation with regard to expanding their search strategy and analysis thought process. We envisage that a balanced sensemaking heuristic would incorporate the strengths and benefits of both the narrow but in-depth and broad but cursory paradigms. The full complement of tools afforded by the RAHS system was designed to support non-prescriptive and flexible workflows, and the team of analysts would do well if they were able to leverage this.

Authors' note: RAHS has undergone major upgrades since the pre-beta version fielded in the October 2006 experiment, and now provides analysts with more capabilities beyond data analysis. Please refer to the paper by Foo et al. (2007).

REFERENCES

Analyst notebook product overview. (2007). i2 Ltd. Products. Retrieved on 3 February 2007 from http://www.i2.co.uk/Products/Analysts_Notebook/default.asp

Anderson, J.R. (1982). Acquisition of cognitive skill. Psychological Review, 89, pp 369-406

Bodnar, J.W. (2005). Making sense of massive data by hypothesis testing. In Proceedings of 2005 International Conference on Intelligence Analysis. Retrieved on February 3, 2007, from https://analysis.mitre.org/proceedings/Final_Papers_Files/124_Camera_Ready_Paper.pdf

Chase, W. G. and Simon, H. A. (1973). The mind's eye in chess. W. G. Chase (Ed.). Visual information processing, pp 215–281. New York: Academic Press.

Chase, W. G. and Ericsson, K. A. (1982). Skill and working memory. G. H. Bower (Ed.). The psychology of learning and motivation 16, pp 1-58. New York: Academic Press.

Ericsson, K.A., Krampe, R. T. and Tesch-Romer, C. (1993). The role of dedicated practice in the acquisition of expert performance. Psychological Review 100, 3, pp 363-406.

Ericsson, K.A. and Lehmann, A. C. (1996). Expert and exceptional performance: evidence on maximal adaptations on task constraints. Annual Review of Psychology 47,pp 273-305.

Foo, K. J., Wong, R. F., Ni, D. N., Leong, M. W. and Leong, H. W. (2007). Developing a horizon scanning system for early warning. Unpublished. Submitted to 12th International Command and Control Research and Technology Symposium.

Klein, G., Phillips, J. K., Rall, E. and Peluso, D. A. (2004). Expertise out of context: Proceedings of the 6th International Conference on Naturalistic Decision Making (Ed, Hoffman, R. R.) Lawrence Erlbaum & Associates, Mahwah, NJ.

Klein, G., Moon, B., and Hoffman, R. R. (2006a). Making sense of sensemaking 1: alternative perspectives. IEEE Intelligent Systems 21, 4 (Jul. 2006), pp 70-73.

Klein, G., Moon, B., and Hoffman, R. R. (2006b). Making sense of sensemaking 2: A macrocognitive model. IEEE Intelligent Systems 21, 5 (Sep. 2006), pp 88-92.

Pirolli, P. and Card, S. (2005). The sensemaking process and leverage points for analyst technology as identified through cognitive task analysis, Proceedings of 2005 International Conference on Intelligence Analysis. Retrieved on 3 February 2007 from https://analysis.mitre.org/proceedings/Final_Papers_Files/206_Camera_Ready_Paper.pdf

Police appeal for assistance - can you help? (2005, July 22). In The UK Police Service – News. Retrieved on 3 February 2007 from http://www.police.uk/content/viewarticle.asp?xslfile=~xsl~transform_article.xsl&xmlfile=~_content~xml~news~MPS-BSW-017.xml

Report of the official account of the bombings in London on 7th July 2005. (2006, May 11). In BBC News. Retrieved on 3 February 2007 from http://news.bbc.co.uk/1/shared/bsp/hi/pdfs/11_05_06_narrative.pdf

Swanson, D. (1987). Two medical literatures that are logically but not bibliographically connected. Journal of American Society for Information Science 38, 4 (1987), pp 228-233

Timeline of the 2005 London bombings. (2007, January 31). In Wikipedia, the Free Encyclopedia. Retrieved on 3 February 2007 from http://en.wikipedia.org/w/index.php?title=Timeline_of_the_2005_London_bombings&oldid=104470337

Thompson, C. (2006, 3 December). Open Source Spying. In The New York Times. Retrieved on 3 February 2007 from http://www.nytimes.com/2006/12/03/magazine/03intelligence.html?ex=1322802000&en=46027e63d79046ce&ei=5090

This paper was first presented at the 12th International Command and Control Research and Technology Symposium in June 2007, Newport, Rhode Island, US, and has been adapted for publication in DSTA Horizons.

BIOGRAPHY

Gwenda Fong is Branch Head (MINDEF Policy Office). She is currently involved in formulating defence relations policies that serve to enhance Singapore's defence and security interests with other countries. Prior to her current appointment, she spent three years at the SAF Centre for Military Experimentation where she was involved in designing, conducting and analysing experiments that explore future operational concepts. She graduated with a Bachelor of Science (Distinction) as well as a Master of Science in Electrical Engineering from Stanford University in 2001 under the Public Service Commission Overseas Merit Scholarship.

Keith Oh is Engineer (Networked Systems). Currently, he is responsible for the design and execution of experiments for the evaluation of emerging technologies for horizon scanning and risk assessment. He graduated with a Bachelor of Science in Electrical & Computer Engineering from Carnegie Mellon University in 2004 and a Master of Science in Human-Computer Interaction from Carnegie Mellon University in 2005 under the DSTA Overseas Scholarship.