



<http://www.diva-portal.org>

Postprint

This is the accepted version of a paper presented at *18th International Conference on Human Interface and the Management of Information (HCI International 2016), Toronto, Canada, July 17-22, 2016..*

Citation for the original published paper:

Helldin, T., Riveiro, M., Pashami, S., Falkman, G., Byttner, S. et al. (2016)

Supporting analytical reasoning: A study from the automotive industry.

In: Sakae Yamamoto (ed.), *Human Interface and the Management of Information: Applications and Services: 18th International Conference, HCI International 2016 Toronto, Canada, July 17-22, 2016. Proceedings, Part II* (pp. 20-31). Springer

Lecture Notes in Computer Science

http://dx.doi.org/10.1007/978-3-319-40397-7_3

N.B. When citing this work, cite the original published paper.

The final publication is available at Springer via http://dx.doi.org/10.1007/978-3-319-40397-7_3

Permanent link to this version:

<http://urn.kb.se/resolve?urn=urn:nbn:se:his:diva-12801>

Supporting Analytical Reasoning

A study from the automotive industry

Tove Helldin¹, Maria Riveiro¹, Sepideh Pashami², Göran Falkman¹,
Stefan Byttner² and Slawomir Nowaczyk²

¹ University of Skövde
`firstname.lastname@his.se`
² Halmstad University
`firstname.lastname@hh.se`

Abstract. In the era of big data, it is imperative to assist the human analyst in the endeavor to find solutions to ill-defined problems, i.e. to “*detect the expected and discover the unexpected*” [22]. To their aid, a plethora of analysis support systems is available to the analysts. However, these support systems often lack visual and interactive features, leaving the analysts with no opportunity to guide, influence and even understand the automatic reasoning performed and the data used. Yet, to be able to appropriately support the analysts in their sense-making process, we must look at this process more closely. In this paper, we present the results from interviews performed together with data analysts from the automotive industry where we have investigated how they handle the data, analyze it and make decisions based on the data, outlining directions for the development of analytical support systems within the area.

Keywords: Analytical reasoning, sense-making, visual analytics, truck data analysis, big data.

1 Introduction

Analysts today are often faced with large amounts of heterogeneous data on which they are to make quick and well-informed decisions. Often, the process of making sense of the data is exploratory, where the ground truth is now known, leaving the analysts in the dark regarding the quality of the reasoning carried out. Moreover, the problems are often ill-defined and the analysis tools used non-transparent. As argued by Stolper, Perer and Gotz [21], the results from the computational reasoning are often presented as an end product that the analysts are to examine, leaving no opportunity for them to guide or inspect the analysis during the process. When dealing with ill-defined problems that are best solved in an experimental manner by a human analyst, the performance of the analyst can be strongly influenced by his/her ability to quickly test different hypotheses, something which is not always possible when dealing with large, streaming,

heterogeneous data. Moreover, by examining the incremental, piecemeal results from the computations, the analyst is enabled to detect uncertainties or abnormalities early on and perform measures to handle them, such as transform the data and adapt the reasoning model(s).

Within the Visual Analytics (VA) research area, efforts have been made to appropriately visualize different types of data and uncertainties and to ensure that the analyst is able to adapt or inspect the reasoning algorithms used. However, the body of the research conducted has been centered around the performance of the computational reasoning strategies, not on how the analyst actually reasons to solve the problem(s) at hand. For example, as argued by Makonin et al. [10], VA has generally not used machine learning techniques within visual interaction to assist and enhance human analytical reasoning. This makes it difficult to adapt the support system in accordance with the analyst's preferences and reasoning style(s) as well as to evaluate the VA systems developed due to our sparse knowledge of how the system assists, if at all, the reasoning carried out.

In this paper, we investigate how a particular group of analysts that handle large quantities of data carry out their analytical tasks and how a VA support system that enables the analysts to interact with the data, the models and the visualizations could aid them in their sense-making process. As a case study, we have interviewed analysts from the automotive industry - a domain where the analysts are in great need of analysis support to handle the large quantities of data involved in order to solve complicated tasks, such as the identification of indicators of the need for different kinds of repairs as well as how the fuel consumption of a vehicle can be decreased. Such predictive analysis can drastically decrease the amount of time that the vehicle has to spend in the workshops, the need for additional spare parts of vehicles, as well as decrease fuel costs and environmental effects.

The paper is structured as follows: section 2 presents a brief review of previous research regarding theories of human sense-making whereas section 3 discusses how the sense-making process can be supported through the use of visual analytics tools. The study is summarized in section 4 and section 5 presents the results obtained. Section 6 offers a discussion of the work conducted, whereas conclusions and directions for future research are presented in section 7.

2 How do experts make sense of data?

This section presents an overview of relevant theories on human analytical reasoning and sense-making. Our aim is to review the literature trying to answer the following question: how do experts analyze data and find insights? We are particularly concerned with the analysis of huge amounts of data from heterogeneous sources and studies that include the analysis of data presented in visual form (we limit our study to individual analysis, collaborative aspects are not included). A more extensive review can be read in [16].

Analysis is cyclic and iterative. Reaching judgment about a single question is normally an iterative process that will produce several more questions about

a larger issue [22]. Depending upon the request's needs, Thomas and Cook [22] distinguish three basic tasks the analysts may be asked to perform: (1) assess, understand the current world around them and explain the past (the product of this type of analysis is an assessment), (2) forecast, estimate future capabilities, threats, vulnerabilities and opportunities and (3) develop options, establish different optional reactions to potential events and assess their effectiveness and implications.

Many forms of intelligence analysis are "sense-making" tasks [14]. Such tasks consist of some kind of information gathering, representation of the information in a schema that aids the analysis, the development of insight through the manipulation of the schema and the creation of some knowledge product or direct action based on the aid (information \rightarrow schema \rightarrow insight \rightarrow product). Sense-making provides a theoretical framework for understanding the analytical reasoning process that an analyst performs. From a psychological perspective, sense-making has been defined as "how people make sense out of their experience in the world" [2]. In [1], the authors describe intelligence analysis as an example of sense-making.

Another framework that explains the analyst's process is the "think-loop model", presented in [1], see Figure 1. The processes and data are arranged by degree of effort and degree of information structure and the data flow shows the transformation from raw information to reportable results. The overall process is organized into two major loops [1]: (1) a foraging loop that involves processes aimed at seeking information, searching and filtering it, and reading and extracting information possibly into some schema and (2) a sense-making loop that involves iterative development of a mental model (a conceptualization) from the schema that best fits the evidence. The analyst integrates [1]: (1) a bottom-up approach that builds a theory based on a hypothesis by assembling evidence assumed relevant to a question, and (2) a top-down approach that searches for evidence for an assumed hypothesis.

The foraging loop is essentially a trade off among three kinds of processes [14]: (1) exploring or monitoring more of the space (increasing the span of new information items into the analysis process), (2) enriching (or narrowing) the set of items that has been collected for analysis and (3) exploiting the items in the set (through reading of documents, extraction of information, generation of inferences, noticing of patterns, etc.). A detailed description of the information foraging theory, models, empirical investigations and applications of the theory to the design of user interfaces can be found in [13].

It is important to note that sense-making does not always follow the progression data \rightarrow information \rightarrow knowledge \rightarrow understanding. For instance, sense-making can have many loops or does not always have a clear beginning or end points (this has been highlighted by Klein, Moon and Hoffman in [9]).

Other researchers have reached similar conclusions about intelligence analysis. Klein, Moon and Hoffman present another view of sense-making in [9], the data/frame theory. For Klein, Moon and Hoffman, a frame is a mental structure that organizes the data and sense-making is the process of fitting information

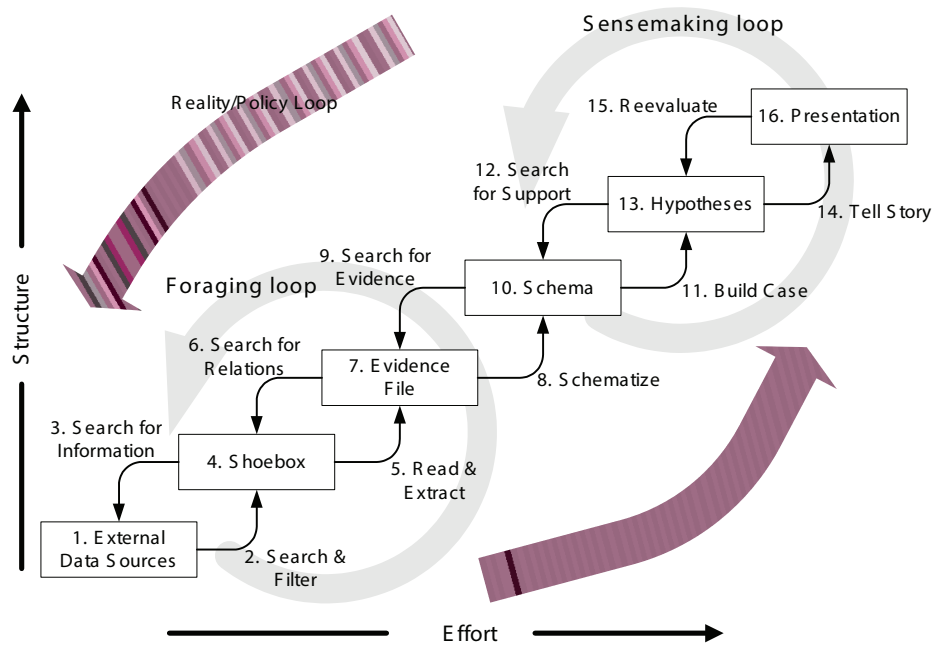


Fig. 1. The think loop model (adapted from [1] and reproduced from [16]). The process is divided into two loops: the foraging loop and sense-making loop.

into the frame. Frames change as data is acquired (frames shape and define the relevant data, and data mandate that frames change in non trivial ways, neither the data nor the frame comes first). Within this theory, sense-making involves two cycles: (1) elaborating a frame and (2) re-framing (questioning the frame and doubting the explanation it provides, leading us to reconsider and seek to replace it with a better one). Frames have similar functionality as the “schema” in the sense-making model by Bodnar [1] and Pirolli and Card [14].

Klein, Moon and Hoffman examine in [9] the meaning of sense-making from three different perspectives: psychology, human centered computing and naturalistic decision-making. Although sense-making and situation awareness have been considered equivalent terms [30], Klein, Moon and Hoffman [9] highlight a main difference between them. Situation awareness (as it is defined by Endsley [3]) is about the knowledge state that is achieved (knowledge of data elements, inferences drawn from the data, or predictions that can be made using these inferences). Sense-making is about the process of achieving these kinds of outcomes, the strategies and the barriers encountered.

Naturalistic decision-making theories explain how experts make complex decisions in real and dynamic environments. A large number of studies in this domain can be considered to describe a sense-making process. Some interesting outcomes for this paper are some assumptions about sense-making that have been proved wrong by naturalistic decision-making research. For example, hu-

man decision makers notice less emergent problems when they passively received automated interpretations [17] (referred as the “data fusion and automated hypothesis generation aid sense-making” myth in [9]). In [12], the authors show that more information does not always lead to better sense-making.

Visual representations are commonly used for data exploration and analysis. Lately, visualization research, and of course, VA is paying a lot of attention to the evaluation of visualizations, and discussions regarding the real value of visualization are filling recent articles. Many authors consider “insight” as one of the main purposes of information visualization, even though they also agree that insight is not a well-understood concept [23] and there is not a commonly accepted definition [11]. In [23], the authors review previous information visualization literature trying to answer “how do people gain insight using visualization?” In their argumentation, they state that sense-making plays a major role in the procedure of gaining insight. Their literature review led to the identification of four types of processes through which people gain insight using visualization: provide overview, adjust, detect pattern and match mental model.

3 Visual analytics for supporting sense-making

Visual analytics has been defined as the science of analytical reasoning supported by visual interfaces [22]. As argued by Green and Maciejewski [4], this analytical reasoning can be divided into two types of analytics: computational analytics, performed by the computerized tools, and reasoning analytics, performed by the human analyst. Computational analytics is suitable for well-defined problems and can often be solved through algorithmic means, whereas reasoning analytics is required when solving ill-formed problems, where the knowledge and creativity of the analyst is vital for finding a solution. Reasoning analytics thus comprises how the analyst interprets the results from the computational analyses, i.e. how the analyst make sense of the data [4] which, in a visual analytics context, is enabled through visualization of and interaction with the computational bases and results. However, the human reasoning strategies involved can significantly vary where the analysts can follow strategies and heuristics such as trial and error, means-end analysis, satisficing, expected utility, abduction, induction and deduction (see [4],[5] for more information).

As argued by Pohl, Smuc and Mayr [15], users of VA systems usually do not develop elaborate strategies in their minds before they start working, but react interactively to what is perceived on the screen. As such, investigating which information to present, how to present it, as well as performing careful analyzes regarding how the VA system user is to interact with the information displayed is of vital importance for guiding the user in his/her individual reasoning process. Important is to consider the limited cognitive capabilities of the analyst to grasp the meaning of the computational analyzes performed as well as to limit the risk of human biases. If overwhelmed with information, the analyst might choose to concentrate on only a limited set of the information available, possibly resulting in that important features are left undetected.

4 Method

To investigate how experts make sense of large amounts of data, we performed interviews with seven data analysts from the automotive industry in Sweden. This particular group of analysts were chosen due to the explosion of data collected from the automotive industry in the recent years and due to the lack of use of visual analytics tools. The purpose of our study was thus to investigate how the analysts work today to make sense of the data collected, as well how the implementation of a VA tool could support this sense-making process in this particular domain (guiding the development of VA tools in this particular domain).

The participants worked at (or cooperated with) two of the largest industries in Sweden, Volvo Trucks and Scania, and had an average of 5,7 years of experience of analyzing data. They were all working with research and development issues within their respective companies, thus focusing on exploratory analyzes of truck data in order to investigate different parameter effects on the driver behavior, causes of vehicle accidents, fuel consumption and maintenance needs of the vehicles. The semi-structured, individual interviews took about an hour to perform and were centered around five themes: data analysis problem statement, data preparation, data analysis, the tools used and their perceived needs for analyzing truck data in an efficient manner in the future. The following section depicts the results obtained from the interviews.

5 Results

5.1 Data analysis problem statement

The interviewees all worked with research and development issues at their respective companies. Thus, the work conducted centered around exploratory analyses of the large amounts of heterogeneous data collected (both from databases and in a streaming fashion) from different sensors on the vehicles, documents from the vehicle workshops, weather and geographical data etc. Their daily work contained tasks such as hypotheses generation, collection of appropriate data sets to answer these hypotheses, data preparation and analysis, and hypotheses verification/rejection. However, one of the analysts argued that what to explore was also data-driven, meaning that also the data, and its quality, could determine which hypotheses that were generated. Three analysts explicitly stated that the data they are working with is quite unique and “one of a kind.” They consider it significantly different from what state-of-the-art methods are typically tested on and developed for. At the same time, as a general rule the data is private and cannot be widely shared – which can be seen as an interesting challenge for the research community.

Every task within this process entails challenges for the analysts. As the work conducted is highly exploratory in its nature, and where there is no ground truth, it is very difficult for the analysts to know how to approach the problem, or if it even is possible. Which variables are likely to have an impact on the hypothesis?

What is the quality of the data that can be used to investigate the hypothesis?
How can the phenomena to be explained best be modeled?

5.2 Data preparation and data analysis

The analysts interviewed had access to several terabytes of data, containing various types of data. None of the analysts interviewed used all the available data in their analyses due to the lack of platforms for handling the amount of data, as well as that several variables in the data were deemed irrelevant for their analysis at hand. Instead, some of the analysts looked for interesting situations in the data, and used such smaller data sets during the analyses. Yet, to find such “interesting situations” could be very time-consuming, such as to find events in video logs of the vehicles without automatic support.

It was interesting to note the difference in opinions between analysts with regards to how big of a problem the necessity of doing initial analysis with a small subset of data was. In some cases this was deemed a serious issue that greatly increased the efforts involved in data analytics, since some results that were promising in such early study would later turn out to be invalid when tested on the whole data set. In other cases, it was considered a normal practice and only a minor burden, as the small subsets of data were found to be representative of the full population. There are several possible explanations for those discrepancies, and it will require further study to determine if they are related to the data analysis process, the task at hand, the experience of individual analysts, the properties of the data, or something else.

There was a noteworthy difference between the analysts working at vehicle manufacturing companies and those working in a professional analytics company. The latter described their main task as adapting the existing data analysis processes to the specific needs of individual customers. In a sense, they saw themselves as experts in data analysis and relied on domain specialists to define the goals. On the other hand, analysts employed in the automotive industry considered themselves more of domain experts. They rely on their own understanding of the subject matter in defining the tasks and evaluating the results.

The analysts in the study had to assess the quality of the data by manual means - often through selecting some small subset of the data and apply quick analyses of it to find outliers, missing values and/or noise. Errors in the data were often due to missing time-stamps and the loss of historical data due to electrical faults. One of the analysts in the study thought of this process as very ad-hoc and argued for the importance of performing continuous and iterative quality investigations of the data used. We also noticed several similarities and differences between the analysts here. Analysts often either said that they had problems with noise or with uncertainty, and did not make a clear distinction between them. We believe the lack of sufficiently good tools for managing big data was the reason as to why the analysis was done by selecting small batches of data instead of using all of the data.

A lot of work is put on pre-processing the data - one of the analysts expressed that *“the data is very, very noisy. Around 90% of my work is about pre-processing*

the data". One obstacle was perceived as the unknown uncertainty in the data. For example, when trying to predict the need for component maintenance, the exact state of the component is not known since the only data available is historical and might include the workshop personnel's subjective options of the state of the component. The lack of ground truth was explicitly mentioned by four out of seven analysts. In addition, the lack of direct interactions with end users or receivers of the analytics results, was also common. This has led to several analysts mentioning that evaluation of results they are obtaining is a challenge.

However, due to the unlimited access to data, one of the analysts argued that noisy data was not a hindrance in the analysis process, since just another data sample set, with better quality, could be used instead. A much greater problem was identified as investigating which variables, or combination of variables, that could have an impact on the hypothesis in focus as well as to create models that could appropriately explain the phenomena in the data. All the analysts were using different statistical, data mining and machine learning methods for regression, classification, clustering, outlier detection, etc.

5.3 Tools and perceived needs for the future

None of the analysts interviewed used tools developed with the framework of visual analytics in mind. The analysts all used standard mathematical and programming tools to detect errors or noise in the data and where they could implement their models and make basic visualizations of the data. The most common tools were Matlab, R and Python, but there were also some who used C++ and Java. Few reported using commercial frameworks and tools such as Hadoop, SPSS or Spotfire and in some cases custom developed tools for their organization. All of the analysts interviewed developed their own scripts and used available tool-kits and libraries to speed up their analyses. In terms of visualizations, the analysts argued that these were only used for two purposes: to better understand the data and its quality, and to understand the computational results. However, one of them argued that he could be significantly aided in his analyses if visualizations could be used to show different views of the data, delimiting the risk for biases and misinterpretations.

When asked about their perceived needs for the future in terms of automatic support that could aid them in their reasoning process, three of the analysts argued for the importance of developing tools and platforms that could handle the increasing amounts of available data, such as data from databases, streaming data and video data. One of the analysts further argued that to be able, during run-time, to make alterations of the analysis performed could significantly decrease the amount of time needed to finish the analyses. Further, to be able to repeat the same pre-processing and analysis steps on a different data set would also decrease the time needed to perform analyses. An interesting observation was that it was common for analysts employed in automotive companies to make a different description of the challenges of the analysis. They tended to focus on that the challenge lies in the uncertain, incomplete or noisy data, i.e. specific properties of the data. The analysts in that worked in a professional analytics

company described a challenge as whether they have the relevant data sources, i.e. a higher level perspective on the analysis.

6 Discussion

6.1 Making sense of data

As stated by Thomas and Cook [22], there are three tasks that an analyst may be asked to perform: analysis to understand the world and the past (assessment), forecasting and action projection. The analysts interviewed dealt with all of these tasks in order to explore their hypothesis. Although similar, the analysts all expressed their analysis process in different ways, highlighting individual differences and perceived challenges when it comes to reasoning. As such, a future VA tool should not only be adaptable to accommodate different types of data, models and tasks, but also to accommodate the possible individual preferences of the users, a fact which has also been highlighted by [8].

The analysts interviewed all tried to make sense of the data in exploratory ways, not knowing the quality of the computational and reasoning analyzes carried out. To aid the analysts in this process, a reasoning tool could present different views of the data used and the results generated in order to, for example, increase the analysts' knowledge of the quality of the data as well as different representations of the results generated. Following the information seeking mantra: overview first, filter and details on demand [20], the analysts could be aided in their sense-making process, offering cognitive support.

To accommodate for the ever increasing amounts of data, support must be provided in order to understand which pre-processing is needed, which parameters are related, as well as to enable the analysts to adjust the computational reasoning carried out during run-time. As such, there is a need for developing effective means of integrating and fusing various types of data (i.e. video data, geographical data, text based data etc.). However, as one analyst concluded, there is a need for developing company internal analysis tools due to data security reasons and the lack of trust in commercial tools.

6.2 Visual and interactive support

The analysts in the study argued that visualization and interaction with the data and computational reasoning results only occurred in the data pre-processing stage when trying to understand the data to be used in the analysis, and in the computational results analysis phase, when trying to understand the results from the automatic reasoning. However, as argued by one of the interviewees, to be able to investigate the progressive computational reasoning during run-time could effectively improve the analysis carried out, especially when dealing with large amounts of data. This interaction capability is the core of the VA framework. However, as argued by Stolper, Perer and Gotz [21], this capability is threatened by the trends of the increasing amounts of data to be used in the

analyses and the development of complex and computational expensive algorithms. These trends forces the user to spend a lot of time waiting to proceed in his/her reasoning process as well as to remember the choices made during the pre-processing and analysis phase in order to understand the results. This challenge has been outlined as one of the major challenges within the VA community, highlighting the need for progressive, incremental visualizations where partial, yet meaningful results from the running analyzes processes are presented to the users [19].

6.3 Future tool

Based on the results from the interviews, we argue that the analysts within the automotive industry could be efficiently supported by developing a reasoning tool enabling the analysts to analyze larger quantities of data, where different views of the data are appropriately presented to detect patterns and outliers and avoid human biases, and where the presentation of the incremental results generated during run-time could enable the analysts to guide the analysis process. Yet, due to the individual differences of the analysts, whose reasoning strategies can be strongly affected by the information presented to them, careful evaluations of such support system must be performed. Due to the implementation of the VA system, the specific user group, the data to be used and the models to be applied, a comprehensive set of evaluation tasks need to be performed in order to investigate if the tool supports the human analyst's reasoning carried out.

Many research papers present user evaluations of already implemented VA applications, such as Jigsaw [7] and CzSaw [6]. However, a few studies propose the execution of evaluative tasks earlier in the design process. For example, Green and Maciejewski [4] have suggested early laboratory studies and in situ evaluations of VA applications, such as field studies, case studies or ethnographics, to capture the reasoning situation of the users as a whole. Through such evaluations, it is possible to investigate how the users interact with the visualizations during the reasoning, as well as to see how reasoning informs the problem solving tasks at hand in the context of use. This view is also supported by Scholtz et al. [18], who argue that the development and evaluation process of VA applications should follow established VA guidelines and user studies, however, where much more work is needed to establish and evaluate easy-to-use and informative VA guidelines.

7 Conclusions and Future Work

In this paper, we have investigated how analysts from the automotive domain analyze large amounts of data in order to solve complicated tasks such as the predictive identification of the need for different kinds of vehicle repairs and how the fuel consumption of the vehicles can be decreased.

The analysts worked at "traditional" companies (i.e. not IT companies), where the tradition of using and analyzing large amounts of data is quite young.

Focus is still on understanding the relationships in the data, and what to make out of it, making the analysis tasks conducted highly exploratory in their nature. The data used is highly uncertain and contains a lot of errors, making the data pre-processing part of the analyses very challenging and time-consuming.

One analyst concluded that the main challenge was not in the methods or tools, but in communicating the results to an immature organization. Due to the recent focus on analyzing data within this domain, there is no tradition of using VA tools for this purpose, but instead traditional analysis tools are used such as Matlab and R. However, a future need for tools that enables the analysts to handle large amounts of data, that can aid the analysts to early detect interesting parts of data sets, patterns and relationships among variables, and different views of the results generated was identified during the interviews.

Future work includes, as Green and Maciejewski [4] suggest, observations of the analysts in their working situation, enabling the extraction of more detailed analysis procedures and difficulties, shedding more light on the complete working situations of the analysts. Additional work further includes the implementation of a big data platform, aiding the analysts to analyze large amounts of heterogeneous data where intermediate results, the models used and the final results can be adjusted and interacted with. A first step could be to implement visual and interactive components of the current tools used and evaluate the impacts on the analysts' analysis performance.

8 Acknowledgment

This research has been conducted within the *A Big Data Analytics Framework for a Smart Society* (BIDAF 2014/32) project, supported by the Swedish Knowledge Foundation. We would like to thank the study participants for their valuable feedback.

References

1. Bodnar, J.W.: Making sense of massive data by hypothesis testing. In: International Conference on Intelligence Analysis. pp. 2–4 (2005)
2. Duffy, M.: Sensemaking in classroom conversations. *Openness in Research: The Tension between Self and Other* pp. 119–132 (1995)
3. Endsley, M.R.: Toward a theory of situation awareness in dynamic systems. *Human Factors: The Journal of the Human Factors and Ergonomics Society* 37(1), 32–64 (1995)
4. Green, T.M., Maciejewski, R.: A role for reasoning in visual analytics. In: Proceedings of the 46th Hawaii International Conference on System Sciences (HICSS). pp. 1495–1504. IEEE (2013)
5. Green, T.M., Wakkary, R., Arias-Hernandez, R.: Expanding the scope: Interaction design perspectives for visual analytics. In: Proceedings of the 44th Hawaii International Conference on System Sciences (HICSS). pp. 1–10. IEEE (2011)
6. Kadivar, N., Chen, V., Dunsmuir, D., Lee, E., Qian, C., Dill, J., Shaw, C., Woodbury, R.: Capturing and supporting the analysis process. In: IEEE Symposium on Visual Analytics Science and Technology (VAST). pp. 131–138. IEEE (2009)

7. Kang, Y.a., Stasko, J.: Examining the use of a visual analytics system for sensemaking tasks: Case studies with domain experts. *IEEE Transactions on Visualization and Computer Graphics* 18(12), 2869–2878 (2012)
8. Keim, D.A., Kohlhammer, J., Ellis, G., Mansmann, F.: *Mastering the Information Age: Solving Problems with Visual Analytics*. Florian Mansmann (2010)
9. Klein, G., Moon, B., Hoffman, R.R.: Making sense of sensemaking 1: Alternative perspectives. *IEEE Intelligent Systems* 21(4), 70–73 (2006)
10. Makonin, S., McVeigh, D., Stuerzlinger, W., Tran, K., Popowich, F.: Mixed-initiative for big data: The intersection of human + visual analytics + prediction. In: *Proceedings of the 49th Hawaii International Conference on System Sciences (HICSS-49)* (2016)
11. North, C.: Toward measuring visualization insight. *Computer Graphics and Applications, IEEE* 26(3), 6–9 (2006)
12. Omodei, M., Wearing, A., McLennan, J., Elliott, G., Clancy, J.: More is better? problems of self-regulation in naturalistic decision making settings. In: Associates, L.E. (ed.) *How Professionals Make Decisions*. pp. 29–42 (2005)
13. Pirolli, P.: *Information Foraging Theory: Adaptive Interaction with Information*. Oxford University Press (2007)
14. Pirolli, P., Card, S.: The sensemaking process and leverage points for analyst technology as identified through cognitive task analysis. In: *Proceedings of International Conference on Intelligence Analysis*. vol. 5, pp. 2–4 (2005)
15. Pohl, M., Smuc, M., Mayr, E.: The user puzzle: Explaining the interaction with visual analytics systems. *IEEE Transactions on Visualization and Computer Graphics* 18(12), 2908–2916 (2012)
16. Riveiro, M., Falkman, G., Ziemke, T., Kronhamn, T.: Reasoning about anomalies: a study of the analytical process of detecting and identifying anomalous behavior in maritime traffic data. In: Tolone, W., Ribarsky, W. (eds.) *SPIE Defense, Security, and Sensing. Visual Analytics for Homeland Defense and Security*. vol. 7346. Orlando, FL, USA (2009)
17. Rudolph, J.: *Into the big muddy and out again: Error persistence and crisis management in the operating room*. Ph.D. thesis, Boston College (2003)
18. Scholtz, J.: Developing guidelines for assessing visual analytics environments. *Information Visualization* 10(3), 212–231 (2011)
19. Schulz, H.J., Angelini, M., Santucci, G., Schumann, H.: An enhanced visualization process model for incremental visualization. *IEEE Transactions on Visualization and Computer Graphics* PP(99), 1–1 (2015)
20. Shneiderman, B.: The eyes have it: A task by data type taxonomy for information visualizations. In: *Proceedings of the IEEE Symposium on Visual Languages*. pp. 336–343. IEEE (1996)
21. Stolper, C.D., Perer, A., Gotz, D.: Progressive visual analytics: User-driven visual exploration of in-progress analytics. *IEEE Transactions on Visualization and Computer Graphics* 20(12), 1653–1662 (2014)
22. Thomas, J., K., C. (eds.): *Illuminating the Path: The Research and Development Agenda for Visual Analytics*. IEEE Computer Society, Los Alamos, CA (2005)
23. Yi, J.S., Kang, Y.a., Stasko, J.T., Jacko, J.A.: Understanding and characterizing insights: How do people gain insights using information visualization? In: *Proceedings of the 2008 Workshop on Beyond Time and Errors: Novel Evaluation Methods for Information Visualization*. p. 4. ACM (2008)