SPIDEQ – A Prototype for Decision Making with Quality Metadata

G. SHANKARANARAYANAN, Babson College (Massachusetts, USA)
BIN ZHU, Oregon State University (Oregon, USA)

Data quality (DQ) metadata is the set of quality measurements associated with the data. Literature has demonstrated that the provision of DQ metadata can improve decision performance. However, it also showed that DQ metadata can overload decision-makers and consequently have a negative impact on decision performance. In this paper, we describe a prototype system, SPIDEQ, for visualizing DQ metadata. We believe that the visualization will shift the data overload from cognitive to perceptual and thus improve the decision capacity of the decision maker. Decision performance will improve even when DQ metadata is provided, a hypothesis that is not addressed in this paper. SPIDEQ is a prototype system that supports the provision of DQ metadata during the decision making process. The results of this study offer insights for the design of decision support systems and the provision of DQ metadata.

Keywords: Data Quality Management, Decision Support Systems, Visualization, Data Quality Metadata, System Prototype, Design Science

1. INTRODUCTION

Research has shown that poor decision quality may be attributed to the poor quality of data used in decision making (e. g., [Redman, 1996], [Ford and Goia, 2000]). Although every organization attempts to improve quality, the costs of reaching near perfect quality typically offsets any benefits gained from very high quality data. Organizations therefore set acceptable quality standards and acknowledge that the data is not of perfect quality [Lee et al., 2006]. Further, improving quality is a continuous process that is achieved over time [Wang, 1998]. Decision-makers must recognize and account for imperfections in the data. It is hence important to communicate data quality to decision-makers.

Yet another reason for communicating data quality to decision-makers is that data quality is contextual [Ballou and Pazer, 2003]. Decisions are made in the context of a particular task, hence data quality must be evaluated contextually. Managerial decision-making tasks are activated by business needs and consist of multiple stages - specifying requirements, gathering information, evaluating alternatives, and formulating decision outcomes [Nutt, 1984; 1998]. The efficiency and success of managerial decision-making are influenced by a number of organizational and individual factors, one of which is the quality of informational inputs [Eisenhardt and Zbaracki, 1992] [Nutt, 1998] [Ford and Gioia, 2000]. And since the quality of the data used in decision-making is an important informational input, understanding data quality is critical for analytical, data-driven decision-making tasks.

From a decision-support perspective, it is important to provide decision-makers with *data quality metadata (DQ metadata)*, *data that describes the quality of the data,* along with the data used. Decision-makers can gauge quality in the context of the task and accordingly, lean more on the better quality data [Shankaranarayanan and Cai, 2006]. Studies have shown that providing DQ metadata can change decision outcomes [Fisher et al., 2003]. Research has shown that providing DQ metadata can improve decision outcomes in structured, data-driven decision-tasks [Shankaranarayanan et al., 2008]. The same research also shows the overload created by the added metadata offsets the positive effect of DQ metadata. Integrating DQ metadata into the decision process requires additional cognitive resources. This overload may leave less cognitive resource for the decision task itself and decision performance may degrade. The positive impact of providing DQ metadata outweighs its negative impact only when decision-makers have the ability to integrate and process DQ metadata. However, research has not examined the mechanisms to provide DQ metadata to decision makers in a manner that permits easier integration of the metadata. Research has shown that visualizing data in graphs can relieve cognitive load by supporting perceptual inferences [Larkin and Simon, 1987]. We posit that visualization can facilitate the integration of DQ metadata into the decision process. Our objective in this research is to describe the design and implementation of a prototype decision support system, SPIDEQ (Special Purpose Interface for Decision Evaluation with Quality metadata) that helps visualize both the data and metadata used in the decision task.

Decision-making can be individual and organizational [Eisenhardt and Zbaracki 1992]. In this study, we focus on individual decision-making. Thompson's model [1967] classifies decision-tasks into four types – analytical, judgmental, bargaining, and inspiration - based on whether the decision-objective and the means to produce results are known. *Our focus is to support analytical decision-tasks in which the means to produce results and the decision-objective are both known.* According to Nutt [1984], analytical tasks often involve large amounts of data to draw inferences. Eisenhardt and Zbaracki [1992] state that analytical tasks belong to rational decisions in which decision makers enter decision situations with known objectives, gather appropriate information, develop a set of alternative actions, and select the optimal one. The process adopted in rational decision making is structured [Keen and Scott Morton, 1978]. Research has shown that a majority of decision tasks in organizations fit this structured decision-making approach [Dean and Sharfman, 1996]. SPIDEQ supports structured, analytical decision-tasks by visualizing DQ metadata.

The remainder of the paper is organized as follows. We first review the relevant literature on data visualization and information representation for decision support. We then describe SPIDEQ's interfaces and how it helps visualize the data and DQ metadata for decision support. We conclude by describing our directions for further research.

2. RELEVANT LITERATURE

2.1 Visualization

To facilitate the integration of data quality metadata into decision-making we need to reduce the cognitive overload. Decision makers will be able to employ a higher quality decision process when the cognitive overload is removed. Consequently, the decision quality will be improved. A successful decision support system can extend the decision makers' "bounds of rationality" [Todd and Benbasat 1999]. Information (or data) visualization plays an important role in it. The term visualization can be broadly defined as "representations of information consisting of spatial, non-arbitrary (i.e. 'picture-like') qualities resembling actual objects or events" [Rieber 1994]. In Information Systems, visualization can be more specifically defined as "*the use of computer-supported, interactive, visual representations of abstract data to amplify cognition*" [Card et al., 1999]. Although data visualization has gained significantly in stature in the context of big data, business intelligence, and analytics, for brevity, we confine our discussions to the basic principles of visualization.

Human beings have separate information processing mechanisms for symbolic information (e.g. language) and graphic information. Dual Coding Theory [Paivio 1991] assumes that memory consists of two separate but interrelated codes for processing information—one specialized for the representation and processing of nonverbal objects/events (i.e., imagery), and the other specialized for dealing with language. Two systems can be activated independently, but there are interconnections between the two that allow dual coding of information. The interconnectedness of the two systems permits cueing from one system to the other, which in turn facilitates the interpretation of context [Rieber, 1994]. Human memory structure theory [Ware, 2000] explains the underlying role of visualization in reducing the cognitive effort of information processing. In the theory, Ware [2000] claimed that human memory structure could be categorized into three layers: iconic memory, working memory and long-term memory. **Iconic memory** is the memory of pre-attentive processing, which extracts multiple visual cues simultaneously. Incoming visual information stays in iconic memory for less than a second before part of it is "read out" into working memory. Visualization is important at this stage because certain visual patterns can be detected at this stage without having to go through the cognition process. **Long-term memory** stores information associated with a lifetime's experiences. **Working memory** integrates information extracted from iconic memory with that loaded from long-term memory for problem solving. The space of working memory is limited. Visualization has the potential to augment the working memory in two ways, *memory extension* and visual *cognition extension*. The high speed loading of external visual input enables visualization to serve as an external memory, thus extending the working memory. Also, visualization can facilitate internal computation by reducing the cognitive load associated with mental reasoning, in other words, cognition extension. Visualization allows decision makers to resort to a perceptual process (requiring less effort) instead of a cognitive process (requiring more effort). These two theories lay the foundation for developing this prototype.

2.2 Data Representation for Decision Support

A key consideration in the design of decision support systems is the format in which data or information is presented [Jarvenpaa, 1989] [Vessey, 1991]. Research has identified presentation as a factor affecting decision performance. [Bettman, 1979] [Tan and Benbasat, 1990][Vessey, 1991]. The representational format for data was an important design feature of SPIDEQ.

Prior research has addressed the different effects of data representation in graphic formats and tables [Jarvenpaa and Dickson, 1988] [Jarvenpaa, 1989] [Tan and Benbasat, 1993] [Vessey, 1991]. In general, graphics lead to better performance for most tasks including summarizing data, showing trends, and comparing patterns. In contrast, tables did a better job in tasks of reading the value of single points [Jarvenpaa and Dickson, 1988]. The advantage of visualizing data in graphic form is that it allows decision makers to resort to a perceptual process instead of a cognitive process when processing data. In most cases, perceptual processes require less effort and consume less time than cognitive process. Under the situation of high cognitive load, perceptual processes facilitate more efficient information processing [Vessey, 1991]. Therefore, the different representations of the problem space *and* possible solutions impact the decision-making/problem solving process and consequently, performance. Data presented in graphical format rather than in table format could help decision makers solve complex problems better [Speier et al. 2003]. When task difficulty increases, graphical forms are more effective than tabular data for representing problem-solving tasks, because decision makers can use visual heuristics to simplify the tasks [Smelcer and Carmel 1997]. As we deal with the increased complexity of the tasks due to the added DQ metadata, SPIDEQ adopts a graphical representation to visualize data and quality metadata.

In addition, Meyer et al. [1999] proposed that graphic display would have an advantage over tables when the displayed data has dependency relationships with other data and when understanding the inter-dependent structure is relevant for the task. Their experimental results supported the advantage of graphic displays over tables when solving the task depended on the understanding of existing dependencies among data. Data quality metadata reveal an important characteristics of the data used in the decision-making tasks. The dependency between data and the corresponding quality metadata represents a type of structure. Moreover, the solutions of the decision-making tasks are influenced by such structures. Therefore, the presentation of both data quality metadata and data in graphic format could benefit the decision-making performance as it would convey the inter-dependent structure to the user. SPIDEQ attempts to do exactly that.

Yet another theory that addresses the data representation is the cognitive fit theory [Vessey 1991]. This uses a similar lens to address the data representation issue arguing that the fit between the problem representation format and the problem solving task can affect the problem solving performance. The basic notion of cognitive fit theory can be described as following: the problem solving process is influenced by both the characteristics of the task and the representation of the problem (different formats of data). A cognitive fit occurs when both the problem-solving task and the representation of the problem space align with the problem solving process. An incongruent fit results in inferior problem solving performance. The cognitive fit theory indicates symbolic representation (e.g. tables) is most appropriate for presenting discrete sets of symbols (e.g. a table of flights departure and arrival times), while graphic formats are most appropriate for showing relationships among discrete sets of symbols [Vessey 1991]. Data quality metadata represent important attributes of corresponding data/information used in the decision-making. The inseparable relationship between quality metadata and data implies that the graphic format to represent data quality metadata along with the data could be more appropriate than presenting data in tables. SPIDEQ attempts to improve the cognitive fit by rendering both the data and DQ metadata simultaneously, improving the representation of the problem and solution space.

3. SPIDEQ PROTOTYPE

The prototype system, SPIDEQ, was designed with the intent of providing the users with an interface that allows users to consume DQ metadata together with the data for decision making in a manner that minimizes/eliminates the cognitive overload. We hence refer to it as a Special Purpose Interface for Decision Evaluation with Quality metadata. While the visualization interface is a big part of SPIDEQ, the prototype also features very easy to use interfaces to capture, manipulate, and evaluate the data and DQ metadata for decision support. SPIDEQ was built using Java with a MySQL backend that maintains each "project" created within SPIDEQ along with its associated data and metadata. The visualization employs the 2-dimensional graphic approach [Shneiderman 1996]. The prototype is designed for structured decision tasks only (many routine decision tasks in organizations are fit in this category). Decision criteria and their relationships represent the structure of the decision task. Users have the flexibility to define and adjust the characteristics of as well as the relationships among various decision criteria (independent variables) to customize each task.

SPIDEQ supports two data input methods. First option is "define the data by user" using which the user may input the required data. Second is "read data from saved files". If option 2 is selected, a file browsing function is activated. This allows users to load data files (the figure corresponding to this functionality is not shown for brevity).

Figure 1 illustrates the SPIDEQ interface that helps define the decision scenario, specifically, the decision criteria and associated rules. The interface has three major components. The panel on the left helps define the decision criteria for a decision scenario. The panel on the left helps define the rules, the relationships and constraints associated with the decision criteria. At this stage in development, the prototype accepts linear relationships only. The central panel is to keep track of the list of criteria which have been defined. To explain the interfaces and the workings of SPIDEQ, let us consider a sample scenario. To evaluate the best DSLR cameras in the market, decision makers typically use five criteria: price, zoom quality (or ratio), photo quality (measured in megapixels), aesthetic appearance, and size of the cameras.

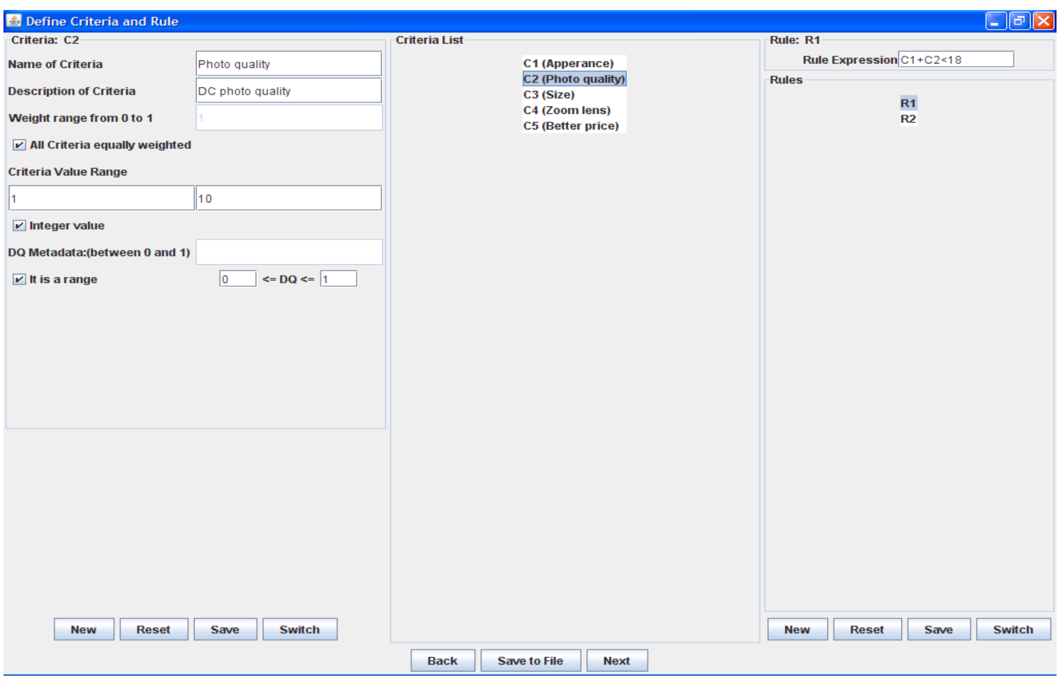

Figure 1: SPIDEQ – Defining Criteria and Rules

To define a new criteria, the user will click on the "New" button. Decision criteria are automatically indexed with names that start with uppercase "C" and a number. For each criteria, users can specify the "Weight" associated with it, if each is to be differently weighted. (The default is that each criteria is weighted the same). In our camera example, the decision maker may choose to weight price and zoom higher than size and appearance. Users can also specify the range of acceptable values for each criteria by defining the upper and lower limits. Returning to our camera example, data on how consumers rate five different brands/models (e.g. Nikon D3300) along each criteria is the data used to make the decision. So, for instance, each brand/model evaluated may be rated on a scale of 1 to 10 for each of the five criteria. Finally, for the data value of each criteria, the DQ metadata (if any) may be specified, either as a range or as a single value. Following literature (e.g., [Pipino et al., 2002], [Lee et al., 2006]), the default is a single value between 0 and 1 with the decimal data type. SPIDEQ also permits the capture of DQ metadata as a range to facilitate "what if" scenarios. Both upper and lower limits accept decimal data between 0 and 1. SPIDEQ is capable of performing basic validation on the data. The newly defined criterion will automatically be added to the "List" in the central panel. To edit a criterion's values, selecting the criterion from the "List" will populate the corresponding values in the left panel. These can then be edited and saved.

The "Switch" button alters the active status between left "Define criteria" and right "Define rules" panel. The "Define rules" panel defines relationships among criteria. "Rules" are also automatically named with names that start with "R". Rules are defined by inputs in the expression field. For example, "R1" specifies that the sum of "C1" and "C2" should be less than 18. The completed rules are listed in the right panel. Selecting each defined rule will allow the user to edit its definition.

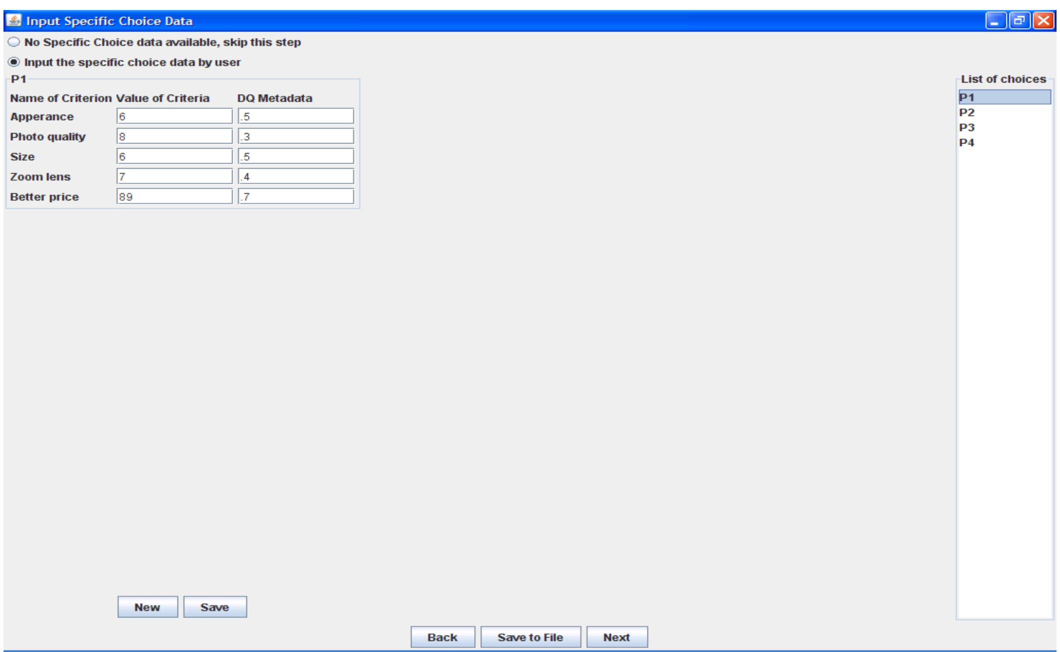

Figure 2: SPIDEQ - Capturing Data on Alternatives

Decision-making often involves choices from a set of alternatives. For instance, the users may be evaluating four different camera brand/models. Let us refer to these as P1 through P4. Each alternative has specific set of values for the decision criteria. Figure 2 shows the interface used to capture such sets of criteria values of those specific alternatives in the decision-making. Returning to our camera example, let PI be a specific brand/model. Figure 2 shows the capture of data values for each of the five criteria defined earlier. Let us further assume that each of these data values is associated with some reliability (DQ metadata) definition. The interface allows the capture of the DQ metadata associated with each data value. SPIDEQ validates the above data against the rules (that relate the criteria) defined and alerts if any of the constraint is violated.

In SPIDEQ, we have adopted the spoke-chart as being the most direct metaphor to visualize both the data and the quality metadata. The spoke chart metaphor is a modification of Kiviat chart [Morris 1974]. Anderson and Dror [2001] adopted the spoke-wheel representation to illustrate multi-objective decision making. SPIDEQ draws from this research for its visualization interface. Each spoke represents a criterion in the decision task. The length of the spoke represents the value of criterion. If the criterion is weighted, then the spoke represents the weighted value – the product of the original value and its corresponding weight. By ensuring that all different scales are standardized to fit in the spoke, SPIDEQ makes it easy to compare visualized displays.

In SPIDEQ, the DQ metadata is shown as a value for each spoke. In addition, a second spoke is superimposed over the first spoke for each criteria. The length of the second spoke is the data value adjusted for its corresponding quality metadata. For instance, if the data value of a specific criterion is of perfect quality, both spokes representing this criterion will be of the same length and will appear as a single spoke. Since quality is rarely perfect, we would see the second spoke overlaying the first. Further, the second spoke is always shorter in length than the first (because the quality will never be greater than 1, will rarely be equal to 1, and in most all cases, be less than 1). Users have the choice of visualizing only the data, only the data adjusted for quality along with the DQ metadata, or both.

To further assist visualization, SPIDEQ allows the spoke representing quality-adjusted data to be color coded. Users can define the color choices. Currently, SPIDEQ has three options - high, medium, and low – to help define colors. The interface also assists the user define the quality range for high, medium, and low to customize how each user gauges the data and its quality in the context of the task. The default values are: 0.8 and higher implies "high" quality data, 0.5 and lower implies "low" quality data, and the "medium" quality is the range in between. The default colors are "green", "yellow" and "red" for the three quality ranges. The interface is shown in figure 3. Such visualization scheme is aligned with visualizing metadata with high uncertainty in the field of Geographic Information Systems [Wise et al. 1995] [Malczewski, 1999]. Based on the guidelines regarding the fit between graphic representation and managerial decision-making proposed by Jarvenpaa and Dickson [1988], grouped bar charts are good fit for presenting how different parts play roles in a whole picture. Cleveland and McGill [1984] have found the human eye to be more accurate in reading grouped bars that have a fixed common baseline. We hence adopted the grouped (or superimposed) spokes for visualization.

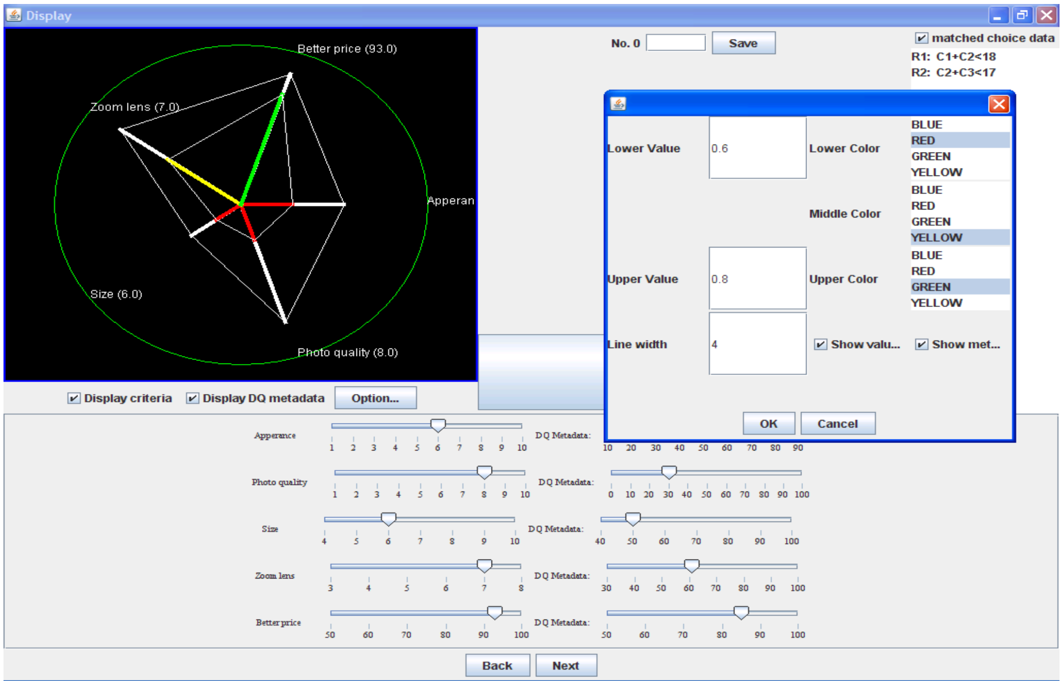

Figure 3: SPIDEQ Interface and Pop-Up Window to define DQ Metadata

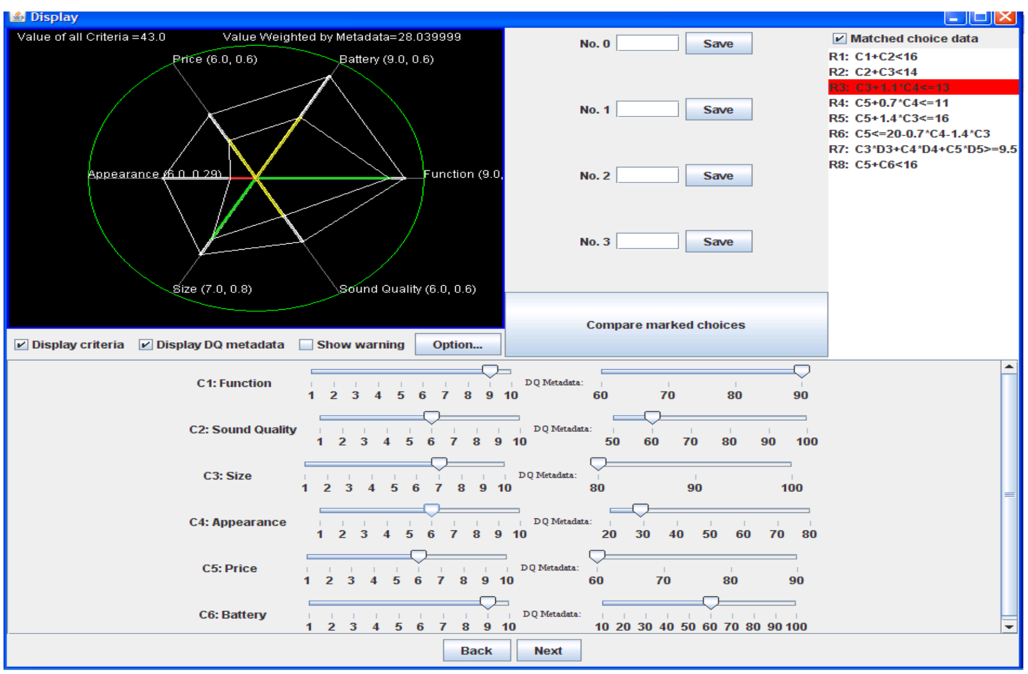

Figure 4: SPIDEQ – Visualizing the decision space.

The outer "web" shows the decision space without DQ metadata. The inner "web" shows the same space with DQ metadata. If the data on each criterion is specified as a range instead of a single value, the sliders at the bottom of the interface in figure 4 may be used to evaluate the alternate scenarios all the values within the range. If any of the rules defined is violated as the user attempts to examine values using the slider, the rule that is violated will appear highlighted in red. The sliders allow the decision-maker to explore the decision space easily and identify viable choices. Due to the inter-relationships, the decision space changes when values of a specific criteria is changed. The "web" changes shape as the decision-maker adjusts values and the visualization allows the decision-maker to see the impact of the changes instantaneously. Figure 4 provides a snap-shot of a point where the decision maker violates a rule during the course of exploring the solution space.

In the upper middle panel, users can mark up to four alternative choices for further comparison. When they use the slider to identify a satisfactory combination of values of various criteria, i.e., a valid solution alternative, users can just click one of the four "Mark choice" slots and name that "choice" with an appropriate name. If the "Compare marked choice" button is pressed, system generates a comparison page displaying up to four named choices.

4. DISCUSSION AND FUTURE STEPS

As elaborated earlier in this paper, the objective of this paper is to describe a prototype system that allows decision makers to incorporate DQ metadata into the decision process. It leverages visualization to shift the cognitive overload due to the added metadata to perceptual, thereby increasing the cognitive capacity of the decision-maker. This addresses a key finding from research: the burden caused by the cognitive overload negatively impacts decision outcome and the decision process.
The interface designs also make the user experience more conducive to effective decision making.

Fisher et al. [2003] posit that decision makers will make better use of DQ metadata if they were trained on how to incorporate it into the decision process. In SPIDEQ prototype, we have hard-coded the method of incorporating DQ metadata using the weighted additive method. While this may be an acceptable method for structured decision making, we are working on improving SPIDEQ to allow decision-makers to interactively customize the manner in which DQ metadata is to be incorporated into the decision process.

We have built SPIDEQ as a first step following the design-science research paradigm in Information Systems research [Hevner et al. 2004]. Adequate information visualization can reduce the cognitive effort of information acquisition, interpretation and integration in decision-making processes. *Build* and *Evaluate* are the two components of design processes identified by March and Smith [1995]. The objective of the second part of this research aims to evaluate whether SPIDEQ will facilitate the integration of data quality metadata into the decision-making process and consequently improve decision performance. We expect to evaluate the following aspects related to the decision process and associated performance.

- Visualizing data allows perceptual processing which typically needs less effort and consumes less time. If data and metadata were presented textually, it is difficult for users to differentiate the metadata and data and to create a mental model of the relationship between the two. We expect the visual representation of data and

DQ metadata to reduce cognitive load in a structured decision environment. Therefore, we expect decision makers with visualization will have lower perceived mental demand in solving decision-tasks compared to decision makers with textual representation.

- If visualization reduces the cognitive demand to integrate DQ metadata into the decision process, decision makers will be able to utilize DQ metadata faster and better. Consequently, all other things being equal, decision makers need not sacrifice decision accuracy in the trade-off between cognitive effort and decision accuracy [Payne et al. 1993]. We expect visualization to improve decision making evidenced by higher objective decision accuracy. We also expect a more efficient decision process evidenced by faster decision time.
- Based on self-efficacy beliefs in Social Cognitive Theory (Bandura, 1986), if a person feels they are capable of achieving the goal, then they are likely to work harder and consequently more likely to succeed compared to a person who is not confident. Therefore, if the presence of DQ metadata impacts the decision makers' confidence, it may indirectly impact decision performance. If a decision maker believes that the DQ metadata provides a more complete picture of the decision task, s/he is likely to have a higher self-efficacy and consequently a higher confidence in the decision. In the absence of a strong theoretical base to support the positive or negative impacts of DQ metadata on subjective decision performance, we expect that decision makers with visualization will be more confident of their decisions compared to the ones using textual representation.

5. CONCLUSIONS

In this paper we have described a prototype decision support tool, SPIDEQ, to help integrate DQ metadata into the decision process. Building on earlier research that posit the negative impact of DQ metadata on decision making due to the increased cognitive load, we have used visualization to reduce this impact. We have described the theoretical rationale for SPIDEQ, covering visualization and data representation for decision tasks. We focus on structured decision tasks and we use a spoke-wheel metaphor to visualize the solution space with and without DQ metadata. We have also briefly described our future research plans based on the design-science research paradigm. We believe this research has important implications for supporting decision-making. The issue of data quality and its importance for decision-making is very evident today in the context of social-media data. Organizations track and mine data from tweets, blogs, and web-based feedback to identify new leads for marketing, increase customer satisfaction, and discover new ideas for product innovation. How do such organizations understand the reliability of the social-media data and decide on the data that is actionable? It is not unreasonable to presume that organizations apply some techniques (e.g., crowd-sourcing) to measure the reliability of social-media data and communicate measurements (rating-values, stars, gamification scores etc.) to decision-makers. Decision-makers range from young, social-media savvy recruits to experienced decision-makers who are not as social-media savvy. This paper presents interesting insights into how organizations should communicate measurements and encourage the use of data and associated metadata (data quality measurements) to such different types of decision-makers.

REFERENCES

Anderson, R. and Dror, M. A. (2001) "Interactive Graphic Presentation for Multiobjective Linear Programming," *Applied Mathematics and Computation*, (123), 2001, pp. 229-248.

Ballou, D.P. and Pazer, H.L. (2003). Modeling completeness versus consistency tradeoffs in information decision systems, *IEEE Transactions on Knowledge and Data Engineering*, 15(1), 240-243.

Bandura A. (1986) *Social Foundations of Thought and Action: A Social Cognitive Theory*. Prentice Hall, Englewood Cliffs, NJ, 1986.

Bettman, J. R. (1979) *An Information Processing Theory of Consumer Choice*, Addison-Wesley Publishing Company, Reading, MA, 1979.

Card, S.K., Mackinlay, J.D. and Shneiderman, B. (Editors). (1999). *Readings in Information Visualization: Using Vision to Think*. Morgan Kaufmann Publishers, Inc., San Francisco

Cleveland, W.S. and McGill, R. (1984). Graphical perception: Theory, experimentation, and application to the development of graphical methods. *Journal of The American Statistical Association*, 79, 531-554.

Dean, J. and Sharfman, M. (1993). Procedural rationality in the strategic decision-making process, *J. Management Stud.,* 30, 587-610.

Eisenhardt, K.M. and Zbaracki, M.J. (1992). Strategic decision-making, *Strategic Management Journal*, 13 (special issue), 17-37.

Fisher, C. W., Chengalur-Smith, I. and Ballou, D.P. (2003). The impact of experience and time on the use of data quality information in decision-making, *Information Systems Research,* 14(2), 170-188.

Ford, C.M and Gioia, D.A. (2000). Factors influencing creativity in the domain of managerial decision-making, *Journal of Management*, 26(4), 705-732.

Hevner, A. R., March, S. T., Park, J. S., and Ram, S. (2004) Design Science in Information Systems Research*, MISQ 28(1),* pp. 75-105.

Jarvenpaa, S. L. (1989). The Effect of Task Demands and Graphical Format on Information Processing Strategies. *Management Science 35 (3) pp. 285-303.*

Jarvenpaa, S. L. and Dickson, G.W. (1988). Graphics and Managerial Decision-making: Research Based Guidelines. *Commun. ACM* 31(6) 764-774.

Keen, P. G. and Scott Morton, M. (1974) *Decision Support Systems: An Organizational Perspective*, Addison-Wesley, Reading, MA, 1974.

Larkin, J. H., and Simon, H. A. (1987) "Why a diagram is (sometimes) worth ten thousand words." *Cognitive Science*, (11), 1987, pp. 65-99.

Lee, Y. W., Pipino, L. L., Funk, J. D. and Wang, R. Y. (2006). *Journey to Data Quality*. MIT Press, Cambridge, MA.

Malczewski, J. (1999). *GIS and Multi-criteria Decision Analysis,* John Wiley & Sons, Hoboken, NJ. (Chapter 7, page 199)

March, S. and Smith, G. (1995). Design and Natural Science Research on Information Technology, *Decision Support Systems, 15, pp. 251-266.*

Meyer, J; Shamo, MK; Gopher, D. (1999). Information structure and the relative efficacy of tables and graphs. *HUMAN FACTORS* 41 (4): 570-587.

Morris, M. F. (1974) "Kiviat Graphs-Conventions and 'figures of merit'", *Performance Evaluation Review,* (3:3), 1974, pp. 2-9.

Nutt, P. C. (1984). Types of Organizational Decision Processing. *Administrative Science Quarterly* 29 414-450.

Nutt, P. C. (1998). How decision makers evaluate alternatives and the influence of complexity, *Management Science,* 44(9), 1148-1166.

Paivio, A. (1991). Dual coding theory: retrospect and current status. Canadian Journal of Psychology 45, 255-87.

Payne, J. W., Bettman, J. R. and Johnson, E. J. (1993).*The adaptive decision maker*. Cambridge University Press, New York, NY.

Pipino, L. L., Lee, Y. W. and Wang, R. Y. (2002). Data quality assessment, *Communications of the ACM,* 45(4), 212-218.

Redman, T. C. (Ed.) (1996). *Data Quality for the Information Age.* Artech House, Boston, MA.

Rieber, L.P. (1994). Computers, graphics and learning. Madison. WI: WCB Brown & Benchmark.

Shneiderman, B. (1996), The eyes have it: A task by data type taxonomy for information visualization, Proceedings of IEEE Workshop on Visual Languages'96, pp. 336-343.

Shankaranarayanan, G. and Cai, Y. (2006). Supporting Data Quality Management in Decision Making, *Decision Support Systems,* 42(1), October 2006, 302-317

Shankaranarayanan, G, Zhu, B., and Cai, Y. (2008) Decision Support with Data Quality Metadata, *Proceedings of the International Conference on Information Quality (ICIQ-2008), Boston, MA*

Smelcer, J. B. and Carmel, E. "The effectiveness of different representations for managerial problem solving: comparing tables and maps." *Decision Sciences* (28: 2) 1997, pp. 391-420.

Speier, C, Vessey, I, Valacich, J. S. 2003. "The effects of interruptions, task complexity, and information presentation on computer-supported decision performance." *Decision Sciences* (34: 4), 2003, pp. 771-787.

Tan, J. and Benbasat, I. 1993. Understanding the Effectiveness of Graphical Presentation for Information Extraction: A Cumulative Experimental Approach, *Decision Sciences*, 24(1) 167-191.

Todd, P. and Benbasat, I. "Evaluating the Impact of DSS, Cognitive Effort, and Incentives on Strategy Selection." *Information Systems Research* (10:4) 1999, pp. 356-374.

Thompson, J.D. (1967). *Organizations in Action*. McGraw-Hill, New York, NY.

Vessey, I. 1991. Cognitive fit-a theory-based analysis of the graphs versus table literature *Decision Science* 22(2) 219-240.

Vessey, I. and Galletta, D. "Cognitive fit: an empirical study of information acquisition". *Information Systems Research*, (2: 1), 1991 pp. 63-85.

Wang, R.Y. (1998). A product perspective on total data quality management, *Communications of the ACM*, 41(2), 58-65.

Ware, C. (2000). *Information visualization: perception for design*, Morgan Kaufmann Publishers Inc., San Francisco, CA.

Wise, J. A., Thomas, J. J., Pennock, K., Lantrip, D., Pottier, M., Schur, A., and Crow, V., (1995) "Visualizing the non-visual: spatial analysis and interaction with information from text documents", in *Proceedings of InfoVis'95, IEEE Symposium on Information Visualization*, New York, 1995, pp. 51-58.

Received Month YYYY; revised Month YYYY; accepted Month YYYY