A MULTI PERSPECTIVE INTER-DISCIPLINARY COMMUNICATION SYSTEM FOR
BUSINESS INTELLIGENCE SYSTEM DESIGN AND DEVELOPMENT

Michael A. Tomasura, Dakota State University, Madison, SD
matomasura@dsu.edu
Surendra Sarnikar, Dakota State University, Madison, SD
surendra.sarnikar@dsu.edu

ABSTRACT
Organizations continue to accumulate large amounts of data as an increasing number of business processes are being conducted electronically. Analyzing these large data sets is often referred to as the “Big Data” problem. While there is an abundance of business data available to business users for analysis, it is not being used due to lack of Business Intelligence (BI) Tools’ capability and a growing backlog of requests to Enterprise IT departments for new and modified data models to support continuously evolving reporting and analysis requirements. In addition, current processes for the design of data models predominantly rely on a sequential and phased approach from requirement collection to data model development. This process is often time consuming and can further exacerbate the large backlogs of requests for new and modified data models at IT departments. This paper addresses the above problem by proposing a collaboration-based tool for use by business users and database developers that can reduce communication gaps and lead to faster development of more comprehensive data models for addressing underlying business needs. The paper presents a discussion of underlying theories such as Interdisciplinary Communication Medium (ICM) and Data/Frame Theory that help understand the problem and examines how they can be applied to develop an IT Artifact to address this issue.

Keywords: Business Intelligence (BI), Data Modeling, Collaboration

INTRODUCTION
Current BI tools and BI development processes lack effective collaboration and communication between database developers and the business users to support the quick and accurate BI development for decision making. This is primarily due to the sequential and phased structure of current processes of data modeling due to which business users typically do not get to see the impact of the changes they request until a prototype is created. According to a recent Forrester Survey (Forrester, 2010) 77% of respondents of a survey report that it takes between days and months to get BI requests fulfilled, 36% of respondents mention that customization is required to answer requests, and 66% of respondents mention too many requests as the main reason for backlog. Easier access to data and data manipulation capabilities for business end users may open up new ways in which business end users may be able to leverage the data for decision making. If database developers cannot provide business users with the needed data models for analysis when they need them, then business users will turn to other options for analyzing data and decision making like spreadsheets. Spreadsheets are often used as a low cost alternative for analyzing data for decision making because of easier access to data manipulation capabilities and
easier access to data through exports. This can lead to many costly problems in the long term in trying to help address the underlying business need.

One of the key objectives of a data modeling and design process is to convey the requirements of business decision making from the business users to a data modeler so the model designed can help address the underlying business need. Difficulties often arise in satisfying this objective since the mental models and BI capability perspectives of business users and database developers vary greatly. Database developers are primarily concerned about optimal data representation and retrieval performance, and these concerns drive their perspectives on data modeling. Business user’s perspectives on data models are driven by the business context of their decision problems and their mental models of the organizational business processes. In order to improve the effectiveness and efficiency of BI development processes, the different logics of data representation and business decision making, and the differences in perspectives on BI capability by database developers and business users need to be taken into consideration during the development process.

The primary objective of this research is to develop a communication and collaboration based tool for BI development process that can help bridge the difference in logics and perspectives among database developers and business end users. Specifically, using the design-science research methodology, a model and a theory-based IT artifact will be developed to address this issue and improve the current BI development process by supporting collaboration between business users and database developers for the design and development of data models. Given that it currently takes days or months to get requirements completed and that requirements change daily to monthly (Forrester, 2010), the key objective of the enhanced multi-perspective collaboration-based BI development process is to reduce the development time and improve the comprehensiveness and quality of requirements capture.

The contributions of this research add to the knowledgebase in two areas. First, the design of the collaboration tool for the BI development process with theoretical foundation in both Interdisciplinary Communication Medium theory and Data/Frame Theory will serve as a template for the development of more user friendly and effective BI development platforms. The design will also add to the knowledge-base of next generation data engineering processes that involve closer interaction with end users. Second, a user observational study of the collaboration-based tool will advance our understanding of how mental models, and views, of data can be communicated more effectively between database developers and business users. It will look at how ambiguity can be removed from the communication of data models requests made by business users.

The remainder of this paper is structured as follows. The next section includes a comprehensive review of relevant work and identifies the research gap that will be addressed in this dissertation. Following the literature review, the theories that form the basis of the IT artifact design are discussed in the theoretical foundations section. The artifact design and implementation section describes the proposed IT artifact design details and the implementation mechanisms, followed by the validation section where observational study plans are described. A summary of the expected research contributions is then presented followed by the future research plan.
LITERATURE REVIEW

Organizations have an overwhelming amount of data in their repositories that needs to be analyzed (J. H. RUSS, 2002; McKinsey, May 2011; Mustaquim, 2011; Uthurusamy, 2002). This data can often exist in their own silos, often with incompatible data structures (Hersh, 2002; McKinsey, May 2011). Some industries have multiple analysis needs like the healthcare field which has two reporting needs: administrative and clinical (T. Mettler, 2008). Putting the two together is not always easy because administrative and clinical data exist in their own silos often with incompatible data structures (Hersh, 2002; Uthurusamy, 2002). Organizations are looking for integration with core information of the business so they can understand what is going on within the organization (Chickowski, 2009). Organizations try to manage the data modeling of large complex datasets using database systems and information management software. Organizations are creating these data models to seek information integration from many large data sources so they can understand what is going on within the organization (Chickowski, 2009).

Getting access to this consolidated data often requires advanced technical expertise and costly time-consuming processes (Kohavi R., 2002). Adding to the demand, a recent study by Forrester Consulting reports that these large data-sets have been growing and have become difficult to analyze using modern database management systems because of the constant change and requests for more data. Also, requirements often change so fast that database developers cannot keep up (Forrester, 2010). Better usable tools for integrating and analyzing data have not been developed (Hersh, 2002; McKinsey, May 2011). Research has called for more user-friendly analytical tools along with a common data representation, with common description language, that will help to make analysis easier (Hans-Peter Kriegel, 2007). However, the data repositories continue to grow and the demand from the business for information continues to grow. A new way to address the demands from the business is needed.

Research has found that BI tools struggle to perform this type of complex analysis and users often do not use the BI tools (Chickowski, 2009). Currently BI tools are often not used because of missing data, incomplete models or because the initial model is no longer relevant. If the data models do not fulfill the needs of the users or they are not intuitive to the users to indicate its capability for analytical requests, it will not be used (Jukic & Nicholas, 2010). User adoption for BI tools has been about 25% since 2005. User adoption of BI tools in 2007 was 25% and 24% in 2009 according to (Howson 2009) and (Swoyer 2010).

BI tools also need to be up to date with the current business views as the business changes and a one-size-fits all approach that is still currently used, is unlikely to satisfy emerging BI needs (Stonebraker, 2005). Different business areas may need to view the data differently and data models need to evolve as quickly as the business does. The issue is that the current tools and change processes lack the collaboration with the business that is needed to develop the models quickly and accurately. According to a recent industry survey (Forrester, 2010), it takes between days and months to get BI requests fulfilled, customization is required to answer requests, business users do not always have a data analyst to create custom reports, and BI requirements change daily to monthly. While user involvement in application development is often limited to the requirements gathering phase, in practice users often have a better understanding of requirements after seeing a prototype (Watson, 2010). However, the long turnaround times for
change requests still needs to be reduced to help organizations get the tools they need to make quicker decisions and tools have not been developed to address this issue.

While several different approaches have been proposed to improve data modeling processes including training (D Batra, 1993; D. Batra, and Wishart, N.A., 2004; Jarvenpaa, 1989), mental aids (D Batra, 1993; D. Batra, 2007; Chen, 1999), restrictive interfaces (Antony, 2005) and knowledge bases, the approaches are predominantly focused on improving the modeling skills of novice data modelers and do not focus on improving user-analyst collaboration to help in the development of more comprehensive and accurate data models that address an underlying business need. Data modeling is often a complex process for business users alone to address even with sophisticated tools and knowledge bases. Given the complexity of data modeling process, collaboration between database experts with knowledge in data modeling, and business users with in-depth knowledge of the underlying business need can be a better approach to the design and development of effective data models for business intelligence.

The specific problem this proposal addresses is the collaboration problem between business users and database developers in regards to data modeling. Research has found that it could take days or months to get requirements completed and that requirements change daily to monthly (Forrester, 2010). With the current process of requesting changes to BI models, the business users do not see what is happening until a prototype is created. Business users often get a better understanding of what they want after seeing the prototype (Watson 2010). Viewing the prototype may lead to scope creep because the business users can now visualize a better solution. Therefore, allowing the business users to watch and participate in the development of models and also having the BI models translated into a business views can help to reduce re-work of the data model development process. Such close collaboration can prevent the development of incorrect or insufficient solutions due to communication issues, different assumptions or different views of the solution, and thus ultimately prevent costly rework and delays in developing an effective BI capability.

RESEARCH METHODOLOGY

In this research, the design science research methodology for information systems is followed for the design, development and evaluation of the proposed information technology artifact (Hevner, 2004). A key contribution of this research is an information technology artifact, the multi perspective inter-disciplinary communication system for BI design and development. This research is relevant to both researchers and practitioners as it addresses an important problem facing organizations in their efforts to leverage organizational data for better decision making and evaluates the effectiveness of inter-disciplinary collaboration systems in supporting design tasks.

The design of the collaboration system is based on an analysis of relevant theoretical foundations in interdisciplinary and collaborative conceptual design frameworks, and data/frame theory that help understand the cognitive processes underlying the BI development process. The utility of the proposed multi perspective inter-disciplinary communication system will be rigorously evaluated with a user observational study. The research contributions of this observational study will advance our understanding of how mental models, or views, of data can be communicated
more effectively between business users and database developers. It will look at how ambiguity can be removed from the communication of data models request made by business users.

This paper tries to address both the technical and managerial audiences that are experiencing issues with data modeling requests as described in the literature review. From a managerial perspective, the contributions of this research can lead to the quicker delivery of data models and reduce the backlog of change requests. Table 1, summarizes the methodology using the seven steps proposed by (Hevner, 2004) in relationship to this research proposal.

Table 1 - Methodology Summary

<table>
<thead>
<tr>
<th>Guideline</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Design as an artifact</td>
<td>This research will result in the development of a model and instantiation of an inter-disciplinary communication system for BI System development.</td>
</tr>
<tr>
<td>Problem relevance</td>
<td>Academic research states that BI tools are not being used because they do not meet users’ needs. Industry research is showing a backlog of request due to long turn over time for IT to deliver new/changing data models. This research addresses the need for collaboration between IT and the business users that will reduce ambiguity in requirements by translating the models between the technical field of IT to the business field.</td>
</tr>
<tr>
<td>Design evaluation</td>
<td>The utility of the proposed multi perspective inter-disciplinary communication system will be rigorously evaluated using a observational study.</td>
</tr>
<tr>
<td>Research Contribution</td>
<td>The contributions will advance our understanding of how views of models and help during design to remove ambiguity in the communication of requirements.</td>
</tr>
<tr>
<td>Research rigor</td>
<td>The model is based on past research and has a theoretical foundation in both Interdisciplinary Communication Medium (ICM) and Data/Frame Theory.</td>
</tr>
<tr>
<td>Design as a search process</td>
<td>The proposed model will be implemented, used and the impacts will be observed. Problems will be identified, solutions will be created and implemented and the cycle can be repeated.</td>
</tr>
<tr>
<td>Communication of research</td>
<td>This paper tries to address both technical and managerial audiences. Management should be interest in this research because it could lead to quicker delivery of data models and reduce the IT backlog.</td>
</tr>
</tbody>
</table>

THEORETICAL FOUNDATIONS

Most design systems follow a sequential mode of design creation that has a predefined workflow of sub-tasks that are sequentially executed and may require numerous iterations making design expensive and time-consuming (Shen, 2008). Prior research that has looked at collaborative development systems promised semi-automated fully integrated tools, networked together, and covering the full product lifecycle (Shen, 2008). More recent research calls for collaborative intelligent user interfaces for human involvement in the design processes (Shen, 2008).
Collaborative systems are used today to assist in communication between IT and Business disciplines; however, none of them try to translate the mental models the users have to assist with the design and development of BI models. One system created by Wang (2003) integrated web-based and agent-based tools to develop a distributed multidisciplinary design optimization environment for collaborative concept design (Wang 2003). This system allowed for interaction between designers, users, and servers but didn't address the conceptual design issues that users experience.

An interdisciplinary collaborate data modeling tool for business users and database developers could help data modelers fulfill BI development requests faster and more accurately by enabling the translation of BI models for the business users as the models are created. There are several theories such as Interdisciplinary Communication Medium (ICM), Data/Frame Theory and Behavioral Decision Theory that can be used to develop an IT Artifact to address this issue.

**Interdisciplinary Communication Medium**

ICM framework is a framework for communication between different disciplines to support collaborative conceptual design and to present a prototype (Fruchter 1996b). This concept suggests that a designer's cycle starts with proposing a shared model, then interprets the model into discipline models with their meaning into the selected discipline context, then critiques the discipline from models to derive behavior and compare it to the functional requirements, and then explains the results to other members of the team (Fruchter 1996b).

The users take a disciplinary approach in the design cycle, it can hinder effective collaboration because of different cultures, educational backgrounds, or design habits of designers (Li, 2006). However, when interdisciplinary communication techniques are used, it reveals differences in the way people think and the way people process data when they are tasked with interdisciplinary work (Winowiecki, 2011). One interdisciplinary communication technique, scenario-building, can be used to help develop interdisciplinary communication (Winowiecki, 2011). This process can be used to help expose conflicts in communication and creates a platform where the conflicts can be communicated and addressed while both parties, in this case the database developer and the business user, trying to develop a BI model.

When the business and IT disciplines have difficulty communicating their mental models they experience delays and miscommunication that may lead to a poor final design and may require rework to correct the issues. The framework of ICM has been found to help other disciplines, such as architectural design or learning environments, to try to communicate designs between different disciplines (Fruchter, 1996a; Winowiecki, 2011). Applying the ICM framework to BI modeling for business users and database developers should help with BI development which also requires a rigorous cross-disciplinary communication of modeling concepts and the decisions used to create the BI models. Database developers need an automated way to convert their BI models to a business view and for business users to convert their views to database developers’ views to help communicate model requirements and lead to reduced time spent reworking in the design process.
There is an opportunity in the early design stage, conceptual design, to have a positive impact on the decisions made to form the models (Wang, 2001). Getting different disciplines to commit to a common view of the models during the design phases has been the most difficult collaborative design task (Wang, 2001). Using the ICM framework with the proposed collaborative system should bridge BI model views, between the two disciplines, required to support the complex design during the design process. The proposed collaborative system will allow each discipline to see the BI models in their own common conceptual view of the actual BI model and an intermediate view that will help with the mental mapping of each discipline to a common view. This collaboration system should lead to reduced rework and less time for the design and development BI models. Frameworks that are used for collaboration and communication that promote participation of business users can contribute to the conceptual design and can lead to an increase in diverse perspectives, higher levels of discourse, and new environments to enhance collaboration (Fischer, 2010).

**Data/Frame Theory of Sensemaking**

Sensemaking is a central cognitive function performed by practitioners in natural settings (Klein, 2007; Weick, 1995). The Data/Frame Theory of Sensemaking suggests that when someone tries to make sense of an event they begin with a perception or a frame (Klein 2006b). The frame concept was originally introduced by Minsky, (1974) who stated that frames defined data. This concept was then extended by (Klein 2006b) as frames themselves actually shape the data and frames change as data is acquired. The Data/Frame Theory assumes sequence between mental model formation and mental simulation (Klein 2006b). The data/frame relationship has been described as something difficult to identify until it is pointed out, after it is pointed out it cannot be missed (Klein, 2007). Past experiences and training help to create people’s frames and therefore form their biases. People sometimes make decisions without even consciously recalling these past experiences. Frames shape data that is measured for Sensemaking and the data itself changes the frame.

The two cycles of Sensemaking are: elaborating a frame and re-framing. Analyzing the frame as data is acquired leads to replacing the frame with a frame that has a better fit for the data. As the frames are refined the data becomes clear. This process of framing and reframing inputs to a problem in a continuous process helps filter and interpret the data (Hutchins, 2011).

In BI development, communication and discovery of the requirements between the database developers and the business users can be framed as a Sensemaking process. The database developers need to understand the business and the business users have to understand consequences of their requests. Business users and database developers can both have different perceptions of the same BI model at the time of design. Data/Frame Theory can be utilized to assist with the reconciliation of the mental models of the database developer frame to the models of the other collaborators data. For example a star schema or dimension hierarchy would be a frame for the database developer. This frame will be reconciled with the business user’s perspective through data.

**ARTIFACT DESIGN AND IMPLEMENTATION**
The proposed novel solution to the collaboration problem between business and IT in regards to BI development is an inter-disciplinary communication system for BI system development that can translate meta-level data in real-time for collaboration between database developers and the business users who are requesting the changes. Table 2 is an overview of the gaps, requirements and the corresponding design contributions of this research. Specifically, we identify a set of critical requirements or objectives that need to be met based on the research gaps. We then explore theoretical frameworks to help understand and address the requirements. On the basis of the supporting theories, we then propose design features of the collaboration system that can help address the requirements.

**Table 2 – Research Framework**

<table>
<thead>
<tr>
<th>Gap</th>
<th>Requirements</th>
<th>Theory</th>
<th>Design Features</th>
</tr>
</thead>
<tbody>
<tr>
<td>- Collaboration problem between business and IT. Business users and data modelers have different mental model.</td>
<td>- Bridge gap of knowledge domains so both business and IT can collaborate and develop data models quicker and more accurately. - Help business users develop requirements to limit scope creep and reduce re-work.</td>
<td>- ICM theory is a framework for communication between different disciplines to support collaborative conceptual design. -Data/frame theory can explain why certain data representations can cause different decisions, between business users and data modelers.</td>
<td>- Interpret metadata content into a business user model (language) as the model changes in real time. - Allow business users to monitor progress of model design and allow them to suggest changes to the model as it is being developed. - Help to remove ambiguity in change request requirements and data modelers interpretation by creating business views containing items that are familiar to business users such as reports, graphs and charts.</td>
</tr>
</tbody>
</table>

Figure 1 shows the key components of the system which should include three views: a business view for the business users, a data view for the database developers creating the BI model view and an intermediate view for collaboration. The business view contains items that are familiar to business users such as reports, graphs and charts that are used to support the decision making process for a business problem. It also includes translation of the data model relationships will be explained in a natural language. The data view contains items that database developers are familiar with such as a star schema with the relationships cardinality and data mappings. The intermediate view contains tables with sample data and annotated diagrams so that both business users and database developers can have a shared representation of the models.
The proposed system uses a web browser connected to the content store, which stores the metadata of the models, along with a form of communication for feedback as the models are being developed. Data models can appear to be complex for business users who do not have a background or training with BI development. As the Database developers change the models the metadata can be translated and displayed in a business user friendly display. Specifically, the business user display consists of data representations that are more familiar to business users such as reports, graphs and natural language representations of semantic data relationships. This view is also web based, and can be accessed from anywhere, with a direct connection to the metadata content data source for real-time model information. Collaboration tools can be used at this point to: work in real-time or be used to mark an object with feedback so when the data modeler is back on-line, the feedback can be addressed.

The new model in the current research would include the items listed in Table 3 Collaboration Tool Requirements. Both business users and database developers could have a different mental model of their perception of what the organization’s BI models should look like and the collaboration system should help to align their perceptions so they could both see the same truth. BI tools use content stores to hold the metadata used for reporting and analysis. To help the business user understand what the database developer is doing, the BI model has to be translated into a model that business users can understand, such a chart or a table. The process of translating the BI model, in real-time, and creating visual displays for the business user should lead to business users being able to clarify their requirements and correct any misinterpretations of the requirements through collaboration. The tool should show improved delivery times for business requests with less rework of the models after the initial requests.

Table 3- Collaboration Tool Requirements

|   | Bridge gap of knowledge domains so both business users and database developers can collaborate. |
2. Interpret metadata content into a business user model (language) as the model changes
3. Help business users develop requirements
4. Allow business users to monitor progress of model design and allow them to suggest changes to the model as it is being developed
5. Help to remove ambiguity between business change request and data modeler

Figure 2 The Business View

Figure 2, the business view, is a visual example of what entities the collaboration view for the business user should contain. The tool will contain the following: a translation into a business language and a visual display of what the output would look like if the design model was approved. There should be a refresh option so the business view can be updated as the database developer makes changes. Finally, there will be a form of communication so questions can be asked and ambiguity can be removed. The communication can be either asynchronous or synchronous. This communication will be recorded and used for analysis and validation.

Figure 3 The Technical View

The technical view, displayed in Figure 3, will contain basic modeling features similar to the commercial modeling tools. The communication tool will also be available in the technical view so feedback can be addressed. The on-line status will help to allow both parties know if the other party is available to talk.
VALIDATION METHOD

In order to validate the proposed system and evaluate whether the research objectives have been satisfied, an observational study will be conducted to test the impact of the proposed tool on the efficiency and effectiveness of the BI development process. A prototype of the multi-perspective interdisciplinary communication system will be created for the observational study. The utility of the proposed interdisciplinary communication system will be evaluated using the user observational study. The observational study will be designed to evaluate the following research objectives: to evaluate if the use of the collaboration system for BI development will reduce rework compared to current sequential BI development process and to evaluate if the quality and amount of time taken to build BI models will be lower when the collaborative tool is used compared to the current sequential BI development process.

The user observational study will include a group of participating students from Dakota State University programs that will be provided with the collaboration tool. About two or three groups of participants will be needed to perform the observation. Groups will consist of one participant from the business discipline and one from the IT discipline. The business discipline will work with the business view and IT discipline will work with the technical view. Both disciplines will work with the intermediate view. The participant’s prior experience with data modeling can vary. The group will be given directions about how the tool works prior to performing the tasks.

The observer would not be in the physical location similar to how social media research is performed (Brown, 2011). Observation data would be taken from interaction logs and used to analyze user’s actions. Participants will compare the existing sequential model for BI development with the collaboration tool. Each participant will start with their own frame based on their own past experiences and training in BI development. The participants will be tasked with a change request that will require data modeling. For example, adding a complex calculation or adding a new entity to an existing model. The tasks will be modeled similar to business requests to simulate a real request. Prior approval from the Dakota State University Institutional Review Board will be requested for this research because it will include human subjects.

We expect the participants will utilize the collaborative system features such as the translation of metadata and the collaborative feedback during the data modeling design phase to have a negative effect on the amount of rework required on the BI development. Rework is measured as 1) the number of wrong requirements or misinterpretation of requirements requiring additional requests and 2) scope creep or the number of new additional requests created after the original request.

The text interaction between participants will be recorded and used for analysis. Interviews will be performed after the tasks are complete to collect additional data and measure the usefulness of the collaboration tools. Open ended questions will be used to encourage a full, meaningful answer using the participants own knowledge and feelings. Example questions will include: Do you feel the collaboration tool was better than the current process of data modeling? Do you feel the multiple views of the model helped you to interpret the models? Why?
Can you give an example of how it helped/hindered the process?

EXPECTED RESEARCH CONTRIBUTIONS

The implications for practitioners is the potential for more accurate BI models that are delivered quicker because less rework would be required in the design and development phases. The contributions of this research include an advancing of the understanding of how mental models, or views, of data can be communicated more effectively between the fields of IT and Business. It will help determine how ambiguity can be removed from the communication of data models requests made by business users and hopefully also into other areas within IS field.

The expected results of the study will be that increased levels of collaboration with translated views of the BI model’s metadata will result in less rework and a shorter design phase and development phase for BI development projects. The contributions of the results will help organizations by providing knowledge about what level of collaboration is needed to help reduce the database development backlog by allowing business users to assist with the design phase. The new collaboration system will address the collaboration problem between business and IT in regards to BI development that is causing the backlogs and delays. It also has potential to prepare organizations for the “Big Data” problem by allowing them to be more agile with the every growing demand and change request.

REFERENCES


Tomasura and Sarnikar

Enhanced Multi-Perspective Collaboration-Based BI Development Process


Tomasura and Sarnikar

Enhanced Multi-Perspective Collaboration-Based BI Development Process


