AN INTERACTION FRAMEWORK FOR BUSINESS INTELLIGENCE

INTRODUCTION

Business intelligence (BI) mainly refers to computer-based techniques used in identifying, extracting, and analyzing business data, such as sales revenue by products and/or departments, or by associated costs and incomes. Common functions of business intelligence technologies are reporting, analytics, dashboard, data exploration and predictive analytics (Moss & Atre, 2003).

With the advances in computer technologies and the widespread of web and mobile applications, improved Users Interfaces (UI) are needed to address different kinds of content consumers. One of the key factors to the success of the Web 2.0 evolution was the emphasis given to the User Experience (UX). Since then, usability and intuitiveness have become a key requirement for any application.

CONTEXT DESCRIPTION

In BI, business users generally perform four types of processes: Data Selection, Data Manipulation, Data Analysis, and Data Visualization. After a full cycle, the user has the possibility to perform these processes again; we call this the BI cycle. Presenting "knowledge-intensive" data for these processes is a common need (Celino & Corcoglioniti, 2010). Interfaces should therefore be visually optimized to present data in the best usable way even in the presence of large data sets. Users need to easily access, view, manipulate, search, extract information from, and discover relationships in their data (Goldstein & Roth, 1994). They typically want to avoid the cognitive load of deeply investigating what they are looking at; they want simple but powerful tools. The simplicity is however somehow missing from all the different BI processes.

Interaction with BI tools such as chart or dashboard reporting tools should take into account many factors. First, the rapidly decreasing costs of network bandwidth, processing power and storage have made BI tools available for a large and diversified set of users. BI tools should therefore be usable by non IT experts. Second, the tools have to deal with different data sources and formats and should hide this heterogeneity to users. Third, BI tools should provide simple interfaces allowing users to create new BI content by mashing up enterprise and external data. Finally, data analysis should be performed keeping in mind the platform's computational capacity. Therefore, capturing the most relevant data and delivering it to the right people at the right time in the right way is essential.

We are living now in the era of mobile computing. We are no longer tied to a browser in a sedentary machine. Users become mobile and will need to work with small devices. This implies that interfaces should communicate results bounded by a limited screen size. This drastic evolution requires new interaction paradigms keeping in mind users` expectations, such as multi touch interaction on tablets, natural gestures and natural language interfaces such as personal assistants.

RESEARCH OBJECTIVES

The objective of this research is to provide a framework that enables users to have a better interaction with the different BI processes by tackling the limitations that affect user experience, and by providing concise and relative recommendation for each process. This will allow users to have better insights and analysis of their data, which will lead to more efficient and accurate decisions. This framework will support the full BI cycle, from data selection, manipulation and transformation to data visualization. Figure 1 depicts the challenges that are raised for each process.

Interactions, Support for Mobility

- Working with Large Data Sets
- Working with External Data Sources using agreed-upon semantics

- Users' End Goal
- Data Exploration
- Users Interactions' Tracing

- Presenting Recommendations, Suggestions and Feedback
- Simplicity

Data Visualization
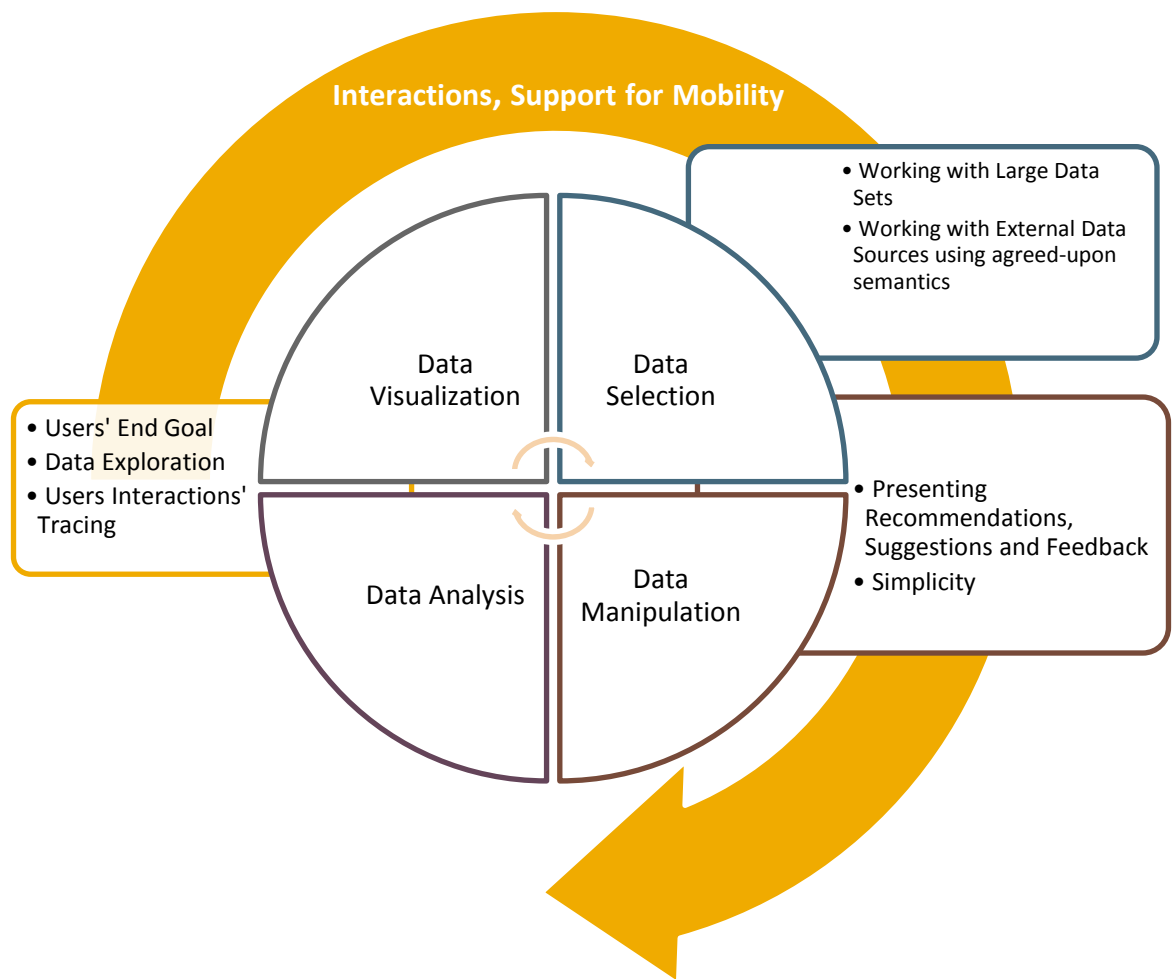
Data Selection

Data Analysis

Data Manipulation

Figure 1 - BI processes and their challenges

The challenges in the Data Selection process stem from working with potentially very large data sets. For example, how can we decide what portion of data to present to users and how can we display this portion of data? Working with external data sources is another challenge. How can we present the different sources? How are we going to design an interface that will present them in an intuitive way?

In the Data Manipulation process a different set of challenges rise. Simplicity in dealing with data is the main focus. How can we design an interface that will intuitively allow users to manipulate data (parsing Values like address strings, parsing and cleaning data/time values, splitting or joining values ... etc.)? Moreover, it is very important at this stage to present the system's suggestions and recommendations that will help guiding users to achieve better results. Collecting feedback and tracing users' interactions are necessary to enhance the quality of future recommendations.

Finally, Data Analysis and Visualization are key elements of this research. They are closely related in the challenges they pose and their importance in the BI cycle. Tracing interaction and analyzing the data presented will help providing more accurate visualization suggestions and guidance. Additionally, if the users express their end goal through the framework, it should be possible to provide them better recommendations and suggestions at each step of the BI cycle. Finally, any Visualization should be able to allow direct interaction and communicate back to users where they are in the processes, where they could go and where they have been. A

set of research questions stem from this requirement especially when dealing with complex visualizations. Some of these questions are:

- How visualizations will be connected to the underlying data?
- Will the user be able to interact directly with the visualization, hiding completely the data behind it?
- How can the framework recommend certain visualizations that match the user's data?

During the different BI processes, there is a continuous interaction from the user's end. A set of several research questions rises when trying to form a guidance framework:

- What are the requirements for designing an intuitive usable interactive interface that will support data selection, manipulation and transformation?
- How can we control the scope of presented data and what are the levels of details needed to decompose and aggregate it?
- How can we identify different user interactions with the framework and what kind of interactions is significant to store?
    - How can we define and present an interaction?
    - How to store this interaction?
    - How to show the set of saved interactions back to the user?
- Can erroneous interactions be captured with the visualization and can suggestions be provided to correct them? How can we decide if any interaction is erroneous or will lead to one that is?

Will it be possible for visualizations to share results and user experiences? This by itself opens a great potential to analyze usage patterns and present recommendations. While interfaces and visualizations interactivity is important, we should always keep in mind its cost. A proper methodology for evaluating interactions needs to be formulated and constant user feedback should be collected. In the end, simplicity is what makes things appealing to users, and this should be the main objective of the framework.

In addition to that, with the vast spread of mobile devices, it is worth investigating how the different findings of the research can impact the user experience on mobile devices and how we can leverage the mobile platform to provide a more intuitive way of interacting with data and visualizations.

RESEARCH METHODOLOGY

The research will progress through the following stages. First there will be an evaluation of the state of the art. This will include a thorough literature survey and a measurement study of current BI systems. The outputs of this phase of research will provide requirements for the next stage which is system design.

Second, an interaction framework will be designed and developed which will address the core issues for the investigation.

Finally, the interaction framework will be evaluated from both a system and user perspective. From the system perspective, requirements for CPU and memory, screens resolution and mobility will be measured and related to the responsiveness of the system. From the user perspective, questionnaires, based on a standard format, will be used to obtain mainly qualitative responses from users. This will be complemented by interviews and focus groups. Interaction logs will be inspected to analyze how users have interacted with the BI tools.

Proposed development:

- Year one: literature review, piloting of tools to be used, development of research techniques, familiarization with data analysis and visualization techniques, development of a small prototype.
- Year two: Main system design and development phase.

- Year three: Analysis of results, supplementary literature review, complete write-up.

BIBLIOGRAPHY

Celino, I., & Corcoglioniti, F. (2010). Towards the Formalization of Interaction Semantics. *I-SEMANTICS* . Graz, Austria: ACM.

Cockburn, A., Karlson, A., & Bederson, B. (2007). A Review of Overview+Detail, Zooming, and Focus+Context Interfaces. *ACM*.

Cyganiak, R., Reynolds, D., & Tennison, J. (2010, July 14). *The RDF Data Cube vocabulary*. Retrieved June 30, 2011, from http://publishing-statistical-data.googlecode.com/svn/trunk/specs/src/main/html/cube.html

Di Lorenzo, G., Hacid, H., Paik, H.-y., & Benatallah, B. (2009). Data Integration in Mashups. *SIGMOD Record*, Vol 38, No. 1, 59-66.

Evelson, B. (2011). *Trends 2011 And Beyond: Business Intelligence.* Forrester Research, Inc.

Goldstein, J., & Roth, S. F. (1994). Using Aggregation and Dynamic Queries for Exploring Large Data Sets. *CHI'94 Human Factors in Computing Systems.* ACM.

Golfarelli, M., Maio, D., & Rizzi, S. (1998). The Dimensional Fact Model: a conceptual model for data warehouses. *International Journal of Cooperative Information Systems*, 215-247 (Volume 7).

Horner, J., Song, I.-Y., & Chen, P. P. (2004). An Analysis of Additivity in OLAP Systems. *Proceedings of the 7th ACM international workshop on Data warehousing and OLAP* (pp. 83-91). Washington, DC, USA: ACM New York, NY, USA.

Jankun-Kelly, T., Ma, K.-L., & Gertz, M. (n.d.). *A Model and Framework for Visualization Exploration.* Mississippi State University, Department of Computer Science and Engineering.

Johnson, J. (2010). *Designing with the - Mind in Mind Simple Guide to Understanding User Interface Design Rules.* E lsevier Inc.

Keefe, D. (2010). Integrating Visualization and Interaction Research to Imporve Scientific Workflows. *IEEE Computer Society*.

Lam, H. (2008). A Framework of Interaction Costs in Information Visualization. *IEEE TRANSACTIONS ON VISUALIZATION AND COMPUTER GRAPHICS,, 14*(6), 1149-1157.

Lam, H., & Munzner, T. (2010). *A Guide to Visual Mutli-Level Interface Design From Synthesis of Emperical Study Evidence.* UBC Computer Science Technical Report.

Landauer, T., & Prabhu, P. (1997). *Handook of Human-Computer Interaction.* Elsevier Science B.V.

Lenz, H.-J., & Shoshani, A. (1997). Summarizability in OLAP and Statistical Data bases. *SSDBM '97 Proceedings of the Ninth International Conference on Scientific and Statistical Database Management* (pp. 132-143). IEEE Computer Society Washington, DC, USA.

Malinowski, E., & Zimanyi, E. (2004). OLAP Hierarchies: A Conceptual Perspective. In E. Malinowski, & E. Zimanyi, *Lecture Notes in Computer Science* (pp. 19-35).

Moss, L., & Atre, S. (2003). *Business Intelligence Roadmap: The Complete Project Lifecycle for Decision-Support Applications.* Addison-Wesley Professional.

O'Reilly, T. (2007). *Google Admits "Data is the Intel Inside".* O'Reilly.

Thor, A., Aumueller, D., & Rahm, E. (2007). Data Integration Support for Mashups. *Proceedings of the Sixth International AAAI Workshop on Information Integration on the Web* (pp. 104-109). Citeseer.