Association for Information Systems
AIS Electronic Library (AISeL)

AMCIS 2009 Proceedings

Americas Conference on Information Systems (AMCIS)

2009

Enhancing Business Intelligence Applications with Value-Driven Feedback and Recommendation

Yoav Kolodner
Ben-Gurion University of the Negev, kolodner@bgu.ac.il

Adir Even
Ben-Gurion University of the Negev, adireven@bgu.ac.il

Enhancing Business Intelligence Applications with Value-Driven Feedback and Recommendation

Yoav Kolodner
Ben-Gurion University of the Negev, Israel
Kolodner@bgu.ac.il

Adir Even
Ben-Gurion University of the Negev, Israel
adireven@bgu.ac.il

ABSTRACT

Business intelligence (BI) systems support activities such as data analysis, managerial decision making, and business-performance measurement. Our research investigates the integration of feedback and recommendation mechanisms (FRM) into BI solutions. We define FRM as textual, visual, and/or graphical cues that are embedded into front-end BI tools and guide the end-user to consider using certain data subsets and analysis forms. Our working hypothesis is that the integration of FRM will improve the usability of BI tools and increase the benefits that end-users and organizations can gain from data resources. Our first research stage focuses on FRM based on assessment of previous usage and the associated value gain. We describe the development of such FRM, and the design of an experiment that will test the usability and the benefits of their integration. Our experiment incorporates value-driven usage metadata - a novel methodology for tracking and communicating the usage of data, linked to a quantitative assessment of the value gained. We describe a high-level architecture for supporting the collection, storage, and presentation of this new metadata form, and a quantitative method for assessing it.

Keywords

Business Intelligence, Recommender Systems, Data Warehouse, Decision Support Systems, Metadata

INTRODUCTION

Data repositories, along with the information systems (IS) that utilize them are critical organizational resources. Recent years have witnessed a transition toward extended use of data for business analysis and decision support, as firms attempt to gain competitive advantage by developing advanced data-analysis capabilities. Our research investigates the integration of *feedback and recommendation mechanisms (FRM)* into business intelligence (BI) system, which support activities such as data analysis, managerial decision making, and business-performance measurement. We define FRM as textual, visual, and/or graphical cues that are embedded into front-end BI tools and guide the end-user to consider using certain data subsets and analysis forms. The working hypothesis of our study is that the integration of FRM into BI tools will improve their usability and increase the benefits that end-users and organizations can gain from data resources.

BI involves acquisition, interpretation, and analysis of data to support managerial decision making. The software market offers a plethora of platforms and tools for supporting BI. Such tools typically offer a variety of presentation capabilities (e.g., tables, charts, statistics, and advanced analytics), and utilities for rapid-development and administrative. Front-end BI tools permit different forms of data usage such as reports, spreadsheets, OLAP (on-line analytical processing), digital dashboards, and data mining. This variety confers the flexibility to use the same data resource for different analytic tasks and to adapt the presentation style to end-users' capabilities and skills. BI system implementations often use a data warehouse (DW) as an infrastructure. The DW stores data covering a broad range of business perspectives and activities over a long time period. In a typical DW, datasets are imported from internal organizational IS, such as enterprise resource planning (ERP) systems (March and Hevner, 2005), and/or from external sources, such as commercial data vendors or the Internet (West, 2000). The imported data are being cleansed, transformed, consolidated, and stored in a repository of historical data about past business behavior, patterns and trends. This back-end DW infrastructure is then used for creating smaller databases (a.k.a., data marts) that can accommodate different analytical needs and thus serve as a platform for supporting BI tasks.

The number of firms engaged in DW/BI implementations and the volumes of data managed have grown immensely in recent years. The increasing popularity of DW/BI can be attributed to benefits such as gaining broad business coverage, leveraging data-collection investments, and shortening implementation cycles (Counihan et al., 2002; March and Hevner, 2005). BI systems enable analytical data usage to support important decisions such as evaluation of corporate strategies and customer segmentation. Yet, exploiting DW/BI environments is challenging both technically, due to the many design decisions involved (Shankaranarayanan and Even, 2004), and organizationally, due to the substantial managerial support and financial resources needed (Wixom and Watson, 2001). Moreover, DW/BI design and configuration decisions are often associated with substantial cost-benefit tradeoffs (Even et al., 2007a). So far, despite the increasing popularity of the DW and the BI concept in recent years, neither has attracted much academic research aimed at the challenge of increasing the effectiveness of DW/BI utilization from the end-user's perspective.

A major limitation of current DW/BI implementation is their complexity from the end-user's viewpoint. The common end-user, in search of an answer to a business question, often finds complex DW repositories difficult to navigate for reaching the right data, and BI tools too difficult to use for answering the question. Furthermore, it is even not uncommon for end-user to neither know the right business question to ask, nor aware of the full range of capabilities offered by DW repositories and BI tools. This limitation exists in current BI solutions, more so with sophisticated interactive tools, which offer advanced visual and analytical capabilities for dynamic and flexible investigation of data, and less so with simple static tools, which offer "snapshot" views, e.g., pre-defined dashboards, reports, or charts. The former tools are geared toward addressing the needs of the data analyst and often require a high level of expertise and an in-depth understanding of the data-resource analyzed, whereas the latter are geared toward supporting the novice user. In terms of economic tradeoffs, sophisticated interactive BI tools offer higher benefit potential, but are costlier in terms of licensing fees and learning curves, whereas simple static BI tools are easier to implement and learn, but offer limited capabilities, and hence, lower benefit potential.

We suggest that FRM capabilities can facilitate effective and efficient navigation by revealing undiscovered potential of unused data and analysis forms, and thus add business value. Next, we present the concept of integrating FRM into BI tools and highlight a few possible approaches for generating them. We focus on FRM based on assessments of previous data usage and the associated value gains. To generate this form of FRM, we propose a novel methodology for tracking the use of data resources, termed as value-driven usage metadata, which integrates in assessments of both the frequency of use and the value gained. We describe architecture for supporting the collection, storage, and presentation of this new metadata form and a quantitative method for assessing it. We then describe the design of an experiment that will test the usability and benefits of FRM integration. To conclude, we highlight the potential contributions of the new concepts that we present – the integration of FRM into BI systems, and the collection of value-driven usage metadata - and discuss directions for future research.

FEEDBACK AND RECOMMENDATION MECHANISMS (FRM)

In this study, we propose to integrate FRM capabilities into BI systems in a manner that would maintain simple and easy-to-learn functionality, while highlighting new usage directions with benefit potential. We define FRM as textual, visual, or graphical cues that are embedded into BI tools, providing the end-user with feedback on the actions that s/he has taken so far, and guiding him/her to consider further actions – e.g., to use certain data subsets and/or to apply certain analysis forms. Such recommendations are common in commercial website for enhancing the end-user's experience (Adomavicius and Tuzhilin ,2005). They can be generated by other users, or by automatic agents, and have been shown to greatly influence end-users' decisions. We suggest that similar enhancement of BI systems may importantly contribute to a better usage of the BI tools, and improve the decisions made.

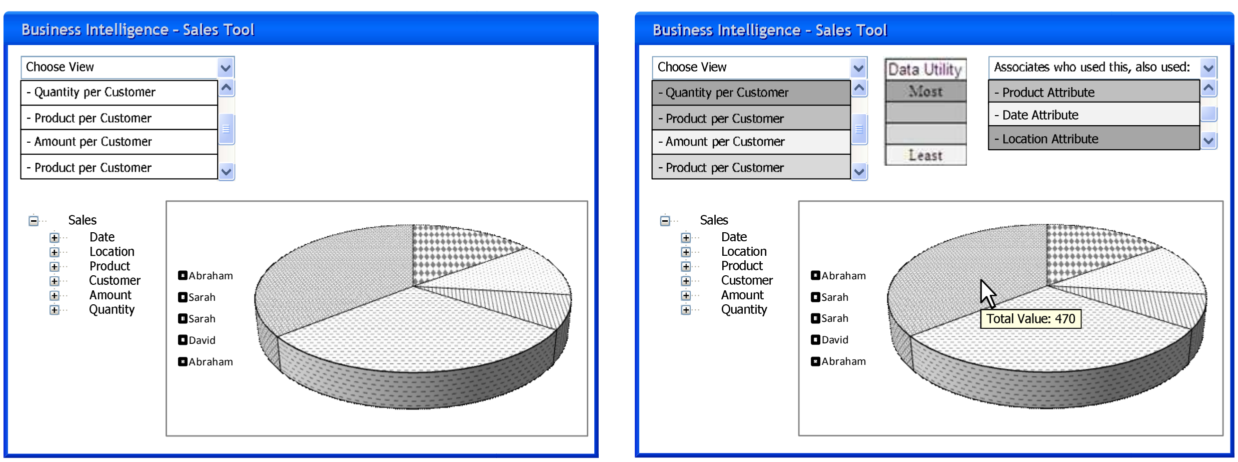

Figure 1. Business-Intelligence Tool (Left) with FRM Added (Right)

Figure 1 offers a simplified illustration of integrating FRM capabilities into a BI tool. The left-hand side BI tool lets the end-user navigate through sales data along certain dimensions (customer, location, date, etc.). This tool treats dimensions equally in the sense that it offers access to all dimensions and leaves navigation decisions to the user. In the FRM-enhanced version of the BI tool (on the right-hand side), navigation decisions are still left to the user, but s/he is provided with some additional visual cues. The cue that Figure 1 demonstrates, for example, is a color-coding that suggests giving higher navigation priority to certain dimensions. Obviously, there are other possible forms for visualizing an FRM besides color-coding, such as textual or graphical pop-up messages and side bars. Such FRM forms could indicate, in addition to the actual recommendations, the level of confidence and relevance of each recommendation based on the parameters that construct it.

We suggest that FRM, when being integrated into BI tools, can facilitate more effective and efficient navigation. This, in turn, may help revealing undiscovered potential of unused data and analysis forms thereby may increase the effectiveness of DW/BI utilization and add business value. In this study we base the FRM on a novel way of collecting metadata- Value-driven usage tracking. Information resources contribute value through usage and experience. In the DW/BI context, the value can be conceptualized, for example, as an objective measure of usage success (e.g., in terms of revenue gained and/or costs saved), or subjectively via an assessment that reflects user satisfaction (Even et al., 2007b). Quantitative assessments of the value associated with the use of data resources have been used to optimize data processes (Ballou et al., 1998) and configure DW datasets (Even et al., 2007a). Even et al. (2007b) highlight the importance of recognizing inequality in the value of data, suggesting that certain data objects, e.g., tables, attributes, and records, may vary significantly in their value contribution. Further, by evaluating a large real-life data resource, the study shows that quantitative assessments of inequality, e.g., by using Gini's index, have important implications for key data management decisions, such as the prioritization of data quality improvement efforts. We suggest that tracking the usage of data and assessing the associated value can be used to construct FRM integrated into BI tools, directing data analysis and exploration and potentially improving decision outcomes – this by providing users with feedback on the outcome of their own usage, as well as an opportunity to benefit from learning how others have gained value from using the same data. Initially, our research will focus on the value-driven approach for generating FRM. We describe a novel methodology for tracking the usage of data resources and linking it to the associated benefits, which we intend to utilize in our first research stage.

VALUE-DRIVEN USAGE METADATA

Data environments are often described as a complex manufacturing process, consisting of interconnected acquisition, processing, storage, retrieval and usage stages (Ballou et al., 1998). These processes can be conceptualized as having two high-level stages – data administration versus data consumption (Figure 2) - each associated with different stakeholders, goals, motivations and tasks. Data administration addresses technical aspects – providing the ICT capacity and performance needed to store and process data resources (e.g., hardware, database systems, and back-end processing), and the tools for implementing information products. Data consumption, on the other hand, would seek to transform information products into business value, through their usage. As the goal of data consumers is typically to gain business benefits and increase profitability, it would be reasonable to assume that they would focus more on the value gained by effective use of data resources and less on the technical aspects associated with managing them.

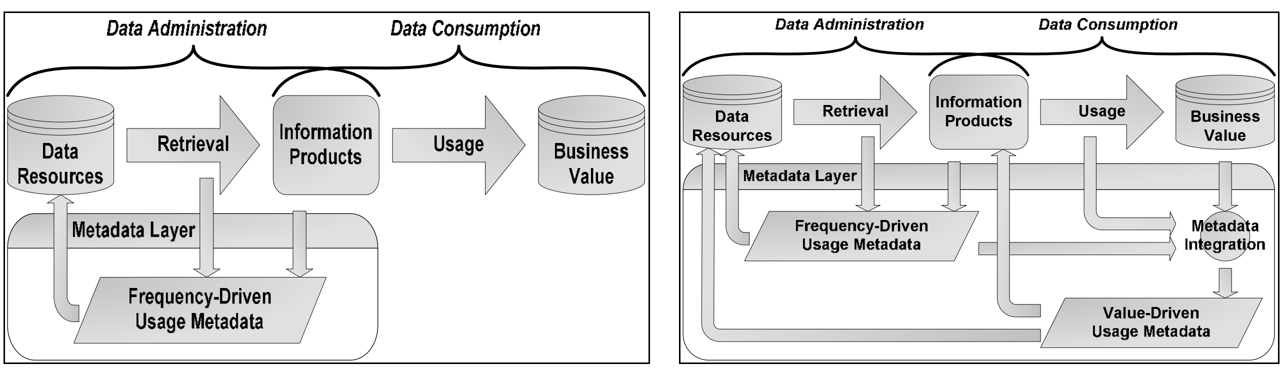

Figure 2. Frequency-Driven (Left) versus Value-Driven (Right) Usage Metadata

Tracking the usage of data objects (e.g., tables, attributes, and records) and applications has been identified as a form of metadata (Shankaranarayanan and Even, 2004). Usage tracking utilities are offered by specialized commercial solutions and, to an extent, by DBMS and BI platforms. We term the common approach implemented by today's solutions as *frequency-driven usage metadata* (the left-hand side of Figure 2) - tracking the queries performed during decision tasks and, by analyzing the queries, identifying the data objects being most-frequently used. The assumption underlying this approach is that frequent usage reflects higher importance. Accordingly, tracking results may affect the configuration and the administration of data resources - e.g., transferring less frequently-used data records to an archive, and/or giving it lower quality-improvement priority (Even et al., 2007a).

To illustrate frequency-driven metadata collection we use the following example (Figure 3). The *Customers* table in this example is used by marketing associates to decide which customers will be approached when promoting a new product. An associate would use a BI tool to investigate previous customer activities, and the tool will generate queries directed to the customers tables, such as those demonstrated, to specify the subset of customers that will be targeted. Each query can be analyzed to detect which records and attributes are used to specify the selection (e.g., by parsing the WHERE clauses in the SQL statement) and, accordingly, the frequency of usage can be determined.

Customers

#	Customer	Gender	Income	Children	Status	Frequency
1	Abraham	Male	High	0	Single	1
2	Sarah	Female	Low	1	Married	2
3	Isaac	Male	Medium	2	Married	1
4	Rebecca	Female	Low	0	Single	1
5	Jacob	Male	Medium	3	Married	1
6	Lea	Female	High	2	Married	3
7	Rachel	Female	Low	4	Single	0
Frequency		3	1	2	1	

Queries

WHERE condition	Attributes Used	Records Retrieved
Gender = 'Male' and Children > 0	Gender, Children	[3], [5]
Gender = 'Female' and Children < 3	Gender, Children	[2], [4], [6]
Gender = 'Female' and Status = 'Married'	Gender, Status	[2], [6]
Income = 'High'	Income	[1], [6]

Figure 3. Illustrative Example (Part 1) - Assessment of Frequency-Driven Usage Metadata

Frequency-driven metadata collection may provide important inputs to the data administrator, toward improving system design and prioritizing administration efforts. It is common in databases, as can be seen in the example above, that some records and attributes are accessed more frequently than others. In a larger real-world databases, this may lead to a decision to give the more frequently used records and attributes higher priority in terms of data quality maintenance – i.e., watch them more closely, detect and correct defects, and make sure to keep them up-to-date.

While seeing the merits of collecting metadata on usage frequency for data administration, we question - does it truly address the needs of data consumers? One could argue that, to an extent, frequent usage reflects higher significance of certain data components from the data consumers' standpoint; hence, a higher value-contribution potential. On the other hand, we suggest that frequent usage may reflect certain stagnation and a tendency to "dig into the same well" - re-using certain data subsets repetitively, while ignoring unused subsets with high contribution potential. Therefore, basing data management decisions on frequency-driven metadata (e.g., transferring less-frequently used data to an archive) has a potential risk - loss of opportunity to benefit from data components that consumers have neglected to use so far, which may permit new forms of data usage.

We agree that there is no "clear cut" answer to this question, as it largely depends on the business context and the usage tasks. However, we suggest that important insights can be gained from tracking and considering not only actual data usage, but also the associated value contribution gains. The benefits gained from the use of information products have been conceptualized as utility; and utility assessments have been used to optimize the configuration of data processes and resources (Ballou et al., 1998; Even et al., 2007). In this study, we suggest that, beyond the benefit offered to data administration, collecting quantitative assessment of the business-value gained as a form of metadata can improve data consumption as well. Business value can be measured in terms of decision outcomes (e.g., production increase, or improved customers' response to promotions), revenues and profitability. Today, such value measurements are captured by organizations, but rarely linked to the data resources and tools that were used to support value generation.

The alternative approach proposed - *value-driven usage metadata* (Figure 2, the right-hand side) - aims at establishing such a link. The baseline for this approach is similar to the collection of frequency-driven metadata – capturing and parsing the queries directed at a data resource. However, with the new approach we also collects different types of value measures (e.g., throughput, performance, and income), and associated with a specific decision task. In certain cases, value assessments can be based on the same data resource analyzed (sale transactions, for example, can often be linked to a specific marketing campaigns based on analyzing previous sales). In other cases – such assessments can be based on other IS resources such as CRM and accounting systems. Value estimates can now be associated with the decision tasks and the queries that supported each tasks, through a mechanism of inference (e.g., by comparing the user-name and the time stamp). Establishing this link between decision tasks and the underlying queries permits the creation of metadata that associate business value with the data components that supported it. Once the value-driven metadata is available, it can be integrated into front-end tools to enhance the presentation, and communicate important information on the frequency of usage and on the associated value to data consumers and administrators. To test this approach, we have successfully implemented a working prototype of a module that captures and stores value-driven usage tracking as part of the metadata layer. In addition, we have developed an API (Application Programming Interface) that can provide this metadata upon demand, through function calls

Once the link between decision tasks and queries is established, different methods can be considered for attributing value to specific data objects. For illustration, we describe here a relatively simple method, which assumes that value is attributed to the last in a sequence of queries that support a decision task. We assume that the task used a tabular dataset with *N* records indexed by *[n]* and *M* attributes indexed by *[m]*. The table is used repetitively to support a certain decision task. We run *Q* queries indexed by *[q]*, each associated with a business value V^q. The binary indicator R^q_n indicates whether record *[n]* was retrieved by query *[q]* ($R^q_n=1$), or not ($R^q_n=0$). Similarly, R^q_m indicates whether attribute *[m]* participated in query *[q]* or not. The value of a certain query *(Vq)* is attributed among the participating data items, using a certain value-attribution function $V^q_{n,m}=u(V^q, R^q_n, R^q_m)$, such that $V^q = \Sigma_n \Sigma_m V^q_{n,m}$. For simplification, we use here an equal attribution of value among all participating data items. Accordingly, the overall value of a certain data items $V_{n.m}$ is given by:

(1) $$V_{n,m} = \sum\nolimits_{q=1..Q} V^q_{n,m} = \sum\nolimits_{q=1..Q} u\left(V^q, R^q_n, R^q_m\right) = \sum\nolimits_{q=1..Q} V^q \Big/ \left(\sum\nolimits_{n=1..N} \sum\nolimits_{m=1..M} R^q_n R^q_m\right),$$ where

Q, q -	The number of queries performed, and the corresponding index, respectively
M, N -	The number of attributes (indexed *[m]*) and records (indexed *[n]*), respectively
V^q, $V^q_{n,m}$, *u* -	The value of query *[q]*, its attribution to data item *[n,m]*, and the function used
R^q_n, R^q_m -	Binary indicators of participation of record *[n]* and attribute *[m]* in query *[q]*

To demonstrate the value allocation described above, we extend the previous example (Figure 4).

#	Customer	Gender	Income	Children	Status	...	Value
1	Abraham	Male	High	0	Single		1000
2	Sarah	Female	Low	1	Married		510
3	Isaac	Male	Medium	2	Married		50
4	Rebecca	Female	Low	0	Single		10
5	Jacob	Male	Medium	3	Married		50
6	Lea	Female	High	2	Married		1510
7	Rachel	Female	Low	4	Single		0
...							
Value		515	2000	60	500		

Value	Color
0	
<100	
<1000	
>=1000	

Queries

WHERE condition	Attributes Used	Records/Value	Total Value
Gender = 'Male' and Children > 0	Gender, Children	[3], [5]	100
Gender = 'Female' and Children < 3	Gender, Children	[2], [4], [6]	30
Gender = 'Female' and Status = 'Married'	Gender, Status	[2], [6]	1000
Income = 'High'	Income	[1], [6]	2000

Figure 4. Illustrative Example (Part 2) - Assessment of Value-Driven Usage Metadata

We assume that each query has led to a certain promotion campaign in which a group of customers has been approached. Some customers may have responded to the campaign by making certain purchases, and the overall value attributed to a query is the total purchase amount. As illustrated, this value proxy may significantly vary among queries. We now use the allocation (Eq. 1) to assess the relative value of each data object. As illustrated by color-coding – some records, and attributes may turn out to have significantly higher value than others, and the value attribution "map" may look significantly different than the one when basing the attribution of usage frequency. We can potentially gain important insights by analyzing the value distribution, along with the assessment of frequency of use. For example, the Income attribute, which was not frequently used, is associated with the highest value, while the Children attribute, which was more frequently used, is associated with lower value. Insights as such can be transformed into valuable recommendations for a marketing associate the next time s/he plans to run a similar campaign. The decision value of each of the participants is being saved in a metadata repository. Using the API that provides access to the value-driven usage metadata - A BI tool, designed to do so, can now demonstrate the value differentials and distribution, as demonstrated in the experimental design described next.

VALUE-DRIVEN USAGE METADATA

This section describes the design of a lab experiment, currently under final preparation stages, that tests the integration of FRM based on previous usage and the associated value.

Research Model and Working Hypotheses

Our first research stage will be directed by the theoretical model shown in Figure 5. Some variables will be measured with the test experiment tool described next, while others will be assessed using a previously-tested questionnaire.

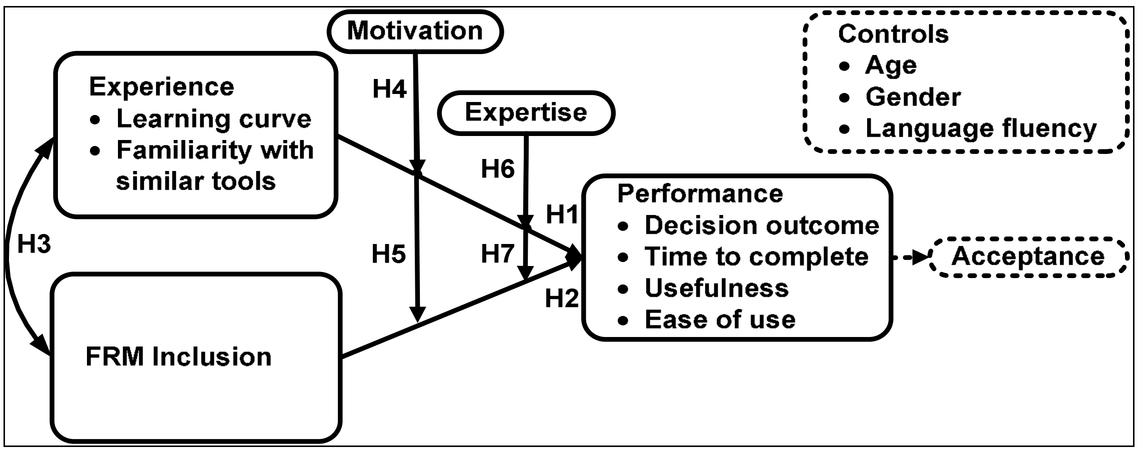

Figure 5. Research Model

Dependent Variable – Performance: The dependent variable, reflecting the ability of a user to effectively perform a task with tool support, will take one of two forms: objective – actual decision outcome and time it takes to complete the decision task, and subjective – perceived usefulness and ease of use. Previously tested models, such as TAM (Technology Acceptance Model), have suggested that a higher sense of usefulness and ease of use increase the likelihood of user acceptance. While this study intends to focus mainly on performance, the experiment described below will permit assessing acceptance and validating the anticipated link between performance and acceptance.

Independent Variables: The independent variables will be *Experience* and *FRM-Inclusion*:

Experience can be measured in terms of learning curve – the time a user spends using and mastering the tool, and familiarity – the extent to which the user has previously used similar tools in the past. It is reasonable to assume that an experienced user (in terms of learning curve and/or familiarity) will perform better that a non-experienced user; hence,

H1: Experience positively affects Performance

Our key assumption is that the inclusion of an FRM will offer a major improvement in the usability of BI tools and therefore in user performance; hence,

H2: FRM inclusion positively affects Performance

As discussed earlier, we suggest that value-driven collection of usage metadata is superior to frequency-driven; hence:

H2a: The Performance effect of value-driven FRM will be superior to the effect of frequency-driven FRM alone

It is reasonable to assume a possible synergistic effect between the two independent variables, i.e., that the overall effect of *Experience* and *FRM-Inclusion* is higher than the effect of each alone; hence,

H3: The interaction effect between Experience and FRM inclusion is positive

Moderating Variables: certain user characteristics may moderate the effect of *Experience* and *FRM-Inclusion* on *Performance*. The moderating variables that will be tested are *Motivation*, the user's motivation to perform well, and *Expertise*, the extent to which the user is knowledgeable in the particular task domain. Studies (e.g., Siegel and Watts-Sussman (2003)) have shown *Motivation* (or involvement) and *Expertise* to have moderating effects on the usefulness of information resources and hence on their acceptance and adoption; hence,

H4: The greater the user's Motivation, the more Experience affects Performance

H5: The greater the user's Motivation, the more FRM inclusion affects Performance

H6: The greater the user's Expertise, the more Experience affects Performance

H7: The greater the user's Expertise, the more FRM inclusion affects Performance

Control Variables: The experiment will control for a few variables - *Age, Gender, Language fluency,* and possibly others.

Experiment Procedures and Tool

The model and the derived hypotheses will be tested in a laboratory setting. In the experiment, all participants will be asked to perform a certain decision task repetitively, aided by a BI tool. The decision outcomes, as well as the actual usage of the tool and the data resources will be tracked and measured, enabling data collection that will allow measuring some of the variables. In addition to tracking decision outcome and actual usage, users will be asked to complete a questionnaire, which will enable data collection on remaining variables. Due to space limitations, we do not describe here in details all the experiment procedures, but rather explain the principles that guide its design.

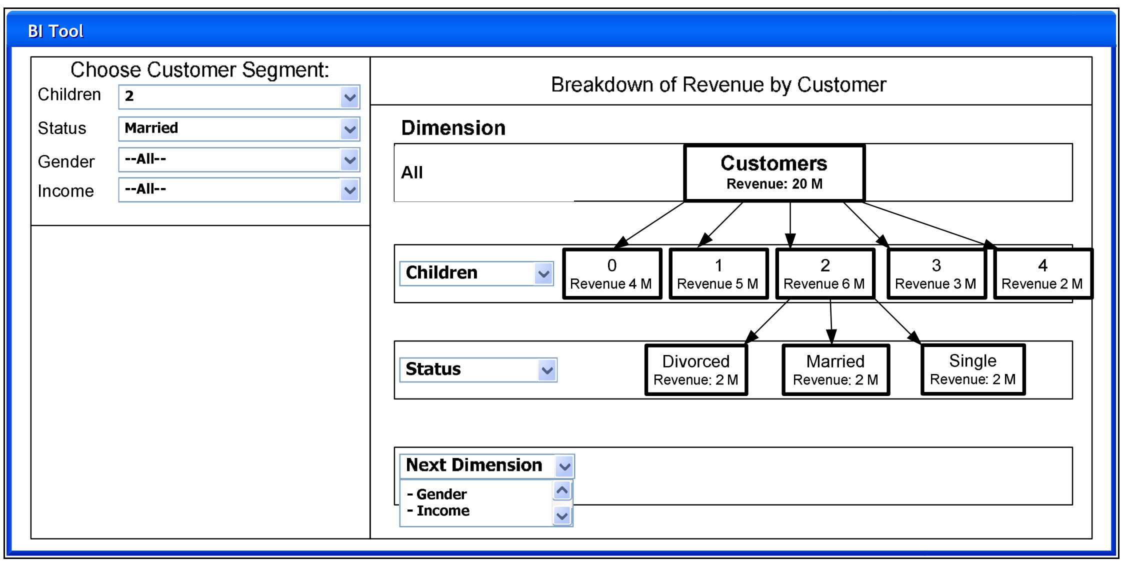

Figure 4. BI Tool Used in the Experiment

Participants in this experiment will act as marketing associates on behalf of a firm that offers a certain service to its customers (e.g., a vacation package). Their decision will be aided by a list of customers, each associated with a given set of characteristics (e.g., Income, Gender, Marital Status, and Number of Children). Based on these characteristics – each customer will be associated with a set of likelihood numbers of purchasing a certain amount units within a given time period. Based on the purchase likelihoods defined per customer, a random generator will produce a large number of transactions for a broad period of time. Aided by a BI tool (Figure 4), the participants will be asked to choose customer segments that will be targeted. Approaching a customer with the promotion offer has a given cost (e.g., the mail delivery fee, or the time needed for a phone call); hence, the larger is the number of customers approached – the higher is the cost. Each customer is associated with certain likelihood to purchase a service, and the overall decision value is defined as the total expected purchase minus the total expected cost. Once the decision is made its value will be calculated and attributed to the different data attributes and records (Eq. 1). The attribution is saved in the value-driven usage metadata module, and as the experiment participants keep performing the decision tasks repetitively – more and more usage metadata is accumulated.

Later in the experiment, a few participants will be provided with an enhanced form of the BI tool (Figure 5), which includes FRM – indication of the total value and the value distribution associated with the different characteristics, at each node. The recommendations are presented to the user on the bottom side of left panel, and change dynamically depending on the node that the user selects. The FRM enhancements are based on the usage metadata that was accumulated while participants keep performing the decision task repetitively.

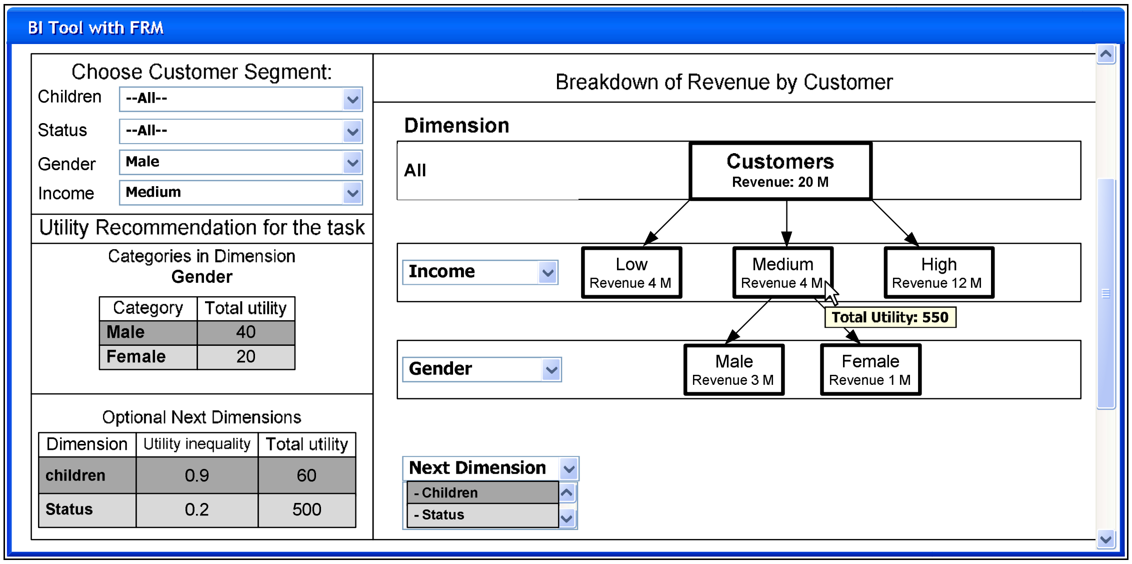

Figure 5. BI Tool Used in the Experiment with FRM Embedded

CONCLUSIONS

Our research investigates the integration of feedback and recommendation mechanisms (FRM) into business intelligence (BI) system. The working hypothesis that guides our study is that the integration of FRM into BI tools will improve their usability and increase the benefits that end-users and organizations can gain from data resources. We have introduced in this study the concept of FRM integration, and described an experiment, currently under preparation, for testing the usability and the benefits of such integration in terms of improving decision-making processes.

Another key contribution of our study, which links to the previous, is the introduction of a novel approach for usage tracking in data environments. This approach suggests that integrating quantitative assessments of usage-frequency together with the associated value gained may offer substantial benefits to data administration and consumption. Joint frequency and value assessments can help identifying unused data subsets with high value-contribution potential, may highlight flaws with repetitive use of data and, consequently, motivate new usage forms. Further, value assessment can direct design decisions, and help prioritizing data maintenance efforts. Relying on usage frequency alone might promote usage stagnation and loss of opportunity to gain new forms of benefits. Complementing frequency assessments with value assessments may help reducing the potential risks. First, value allocation gives higher weight to past usages with high contributions. Second, it can reflect variability in the importance of different subsets depending on the usage context. Lastly, it can help detecting data subsets with high contribution potential that have not been frequently used.

Obviously, future extensions to our study will need to address key limitations: *(a) Quantifying value* – organization maintain performance measurements (e.g., productivity, income, and profitability) that can be linked to decision tasks. However, decision value may depend on the data-usage context and on other resources such as human knowledge and financial assets. *(b) Linking value to specific queries* - performance assessments are rarely linked explicitly to data resources. Our prototype includes inference mechanisms for creating such links. Obviously, implicit links cannot be absolutely precise and might bias the value allocation significantly. Establishing explicit links will require stronger metadata integration between systems and, likely, redesign of data environments (e.g., joint codes that link each decision task and queries), and *(c) Attributing value to specific data objects* – the attribution system has critical impact on the results. Our prototype attributes value only to the last query in the sequence that generated the decision, and distributes the value equally between all the data that were retrieved. A different allocation method may consider, for example, attributing the value to a sequence of queries and/or consider possible interactions among attributes.

ACKNOWLEDGEMENT

This project was supported by a grant from Microsoft R&D, Israel.

REFERENCES

1. Adomavicius, G., Tuzhilin A (2005), Toward the Next Generation of Recommender Systems: A Survey of the State-of-the-Art and Possible Extensions, IEEE Transactions on Knowledge and Data Engineering, 17 (6). 734-749
2. Ballou D. P., Wang R., Pazer H., and Tayi G. K. (1998) Modeling information manufacturing systems to determine information product quality, *Management Science*, 44, 4, 462-484.
3. Counihan, A., Finnegan, P., and Sammon, D. (2002) Towards a framework for evaluating investments in data warehousing, *Information Systems Journal,* 12, 321-338
4. Even, A., Shankaranarayanan, G., and Berger, P. D. (2007a). Economics-driven design for data management: an application to the design of tabular datasets, *IEEE Transactions on Knowledge and Data Engineering*, 19, 6, 818-831
5. Even, A., Shankaranarayanan, G., and Berger, P. D. (2007b). Inequality in the utility of data and its implications for data management. *The Seventeenth Annual Workshop on Information Technologies and Systems (WITS)*, Dec. 2007, Montreal, Quebec, Canada
6. March, S. T. and Hevner A. R. (2005) Integrated decision support systems: a data warehousing perspective. *Decision Support Systems,* 43, 3, 1031-1043
7. Shankaranarayanan, G., and Even, A. (2004) Managing metadata in data warehouses: pitfalls and possibilities. *Communications of the AIS 2004*, 14, 247-274
8. Watts-Sussman, S., and Siegel, W. S. (2003), Informational influence in organizations: an integrated approach to knowledge adoption. *Information Systems Research.* 14 , 1, 47-65
9. West L.A. Jr. (2000) Private markets for public goods: pricing strategies of online database vendors. *Journal of Management Information Systems*, 17, 1, 59-84
10. Wixom, B.H., and Watson, H.J. (2001) An empirical investigation of the factors affecting data warehousing success. *MIS Quarterly*, 25, 1, 17-41