PRACE NAUKOWE AKADEMII EKONOMICZNEJ WE WROCŁAWIU
Nr 1134 2006
Nowoczesne technologie informacyjne w zarządzaniu

Moh'd Al-Sqour, Kamal Matouk

A BUSINESS SCENARIO FOR LEARNING ANALYTICS DATA WAREHOUSING

1. Introduction

Growing volumes of discrete sets of data is forcing management to use technology as a proactive enabler for managing client relationships. In many companies, the critical business data are locked inside transactional systems, and other disparate file sources [Inmon 2002]. These data must be consolidated to have supporting metrics for planning activities. The business leaders utilize systems that integrate all aspects of their corporate information reporting to deliver insight to decision makers. Here non-analytical reporting applications is static, and largely disconnected from other enterprise data that would allow to identify trends and correlations between learning and organizational objective achievements. Therefore, a successful system to meet the above needs cannot be satisfied by simply using a relational database model approach. While well suited to managing transactions, in terms of extensibility and storing large amounts of data, relational databases are typically unable to handle ad-hoc, speed of thought analytical querying and sophisticating business calculations for large user communities. Non-analytical reporting applications provide high value when it comes to static regular operational reports. However, these reports are static, and largely disconnected from other enterprise data that would allow you to identify trends and correlations between learning and organizational objective achievements. More specifically, reports are largely pre-defined and limited to their original intent.

A DW can serve as a storage medium for keeping the corporate memory, or at least concerning certain types of data [Gonzales 2003]. It helps gaining new knowledge by delivering well integrated data to analysis tools, and thus becomes an important part of Decision Support Systems (DSS). In this way a DW, storing only data, results in growth of knowledge and may lead to enhance the enterprise's

success. DW contains several kinds of data which are accessed through analysis front ends, such as OLAP. OLAP technology represents core mechanisms exploited within DSS and it is through this system that management decisions are based on assured, enterprise wide, and real data. Analytics software has emerged recently as a special offering within the range of computer-based business intelligence (BI) tools. BI systems are broad in scope and feature a range of analytical and reporting functionalities. In contrast, analytical tools are easier to use, facilitating quick searches of comprehensive databases for both market, and customer – related data. Traditional reporting systems tell you what happened, with a learning analytics system, you should not only know what happened, you should also know why it happened, and what action you can take to create continuous improvement in your organization. A learning analytics system provides immediate access to the information to improve training operations, reduce costs, better negotiate vendor agreements, and improve business impact. More vitally, this information should be in a form that is clearly and credibly valuable to line of business managers and executives who now have the link they need between learning and ultimate organizational success.

An effective learning analytics system must act as the conduit to information that is otherwise hidden from view, and be so user-friendly as to be an instant addition to the decision-support infrastructure of the learning program. The key attributes of a learning analytics system must be: User-friendly, out of the box analyses, efficient data-sharing capabilities, and Architecture scalable to the needs of the enterprise. A learning analytics system is intended to be a different solution to commonly used operational reporting applications and tools. With a learning analytics system, you should have immediate access to business analyses, business metrics, and a flexible interface to group and filter data; drill up and down through charts. Learning analytics system should be built around subject areas, making it easy to find the answer to common business questions rapidly. Each subject area, in turn, should include a set of pre-defined analyses, designed to answer a wide range of business questions. Within each subject area, users would therefore view a standard analysis, drill up and down, change dimensions and filters, and, where selected, chart information. While it is impossible to project all the analysis needs of every learning organization.

To make analyses more powerful, you should be able to create subscriptions so that your line of business managers, directors, and executives can monitor operations directly from their email inbox. These subscriptions should arrive with the intended user by email or other natural delivery medium, and must be sufficiently easy to comprehend as to require no upfront training. Users who receive subscriptions should not need to be users of the learning analytics system, and the subscribed and delivered analyses should be viewable offline.

2. Implementing Learning Analytics

A DW is a central repository for all or significant parts of the data that an enterprise's business systems collect. Information from various learning and business systems is extracted and organized within the DW for use by analytical applications and user queries. In learning systems, as transactions occur on a day-to-day basis transactions are sent to the learning data warehouse for consolidation and subsequent analysis. The data sent to the warehouse should contain attributes for the various dimensions. In many cases, the dimension attributes will be part of the original transaction. In other cases, the computer programs that are sending information to the warehouse will need to collect and format the dimension attributes prior to sending the transactions along. When the data arrive in the warehouse, they are organized according to the model so that the transactions can be aggregated into measures and accessed across the various dimensions. The DW only provides tools for selecting and displaying statistics. It is the job of the analytic software to provide the presentation interface and the algorithms to aggregate, compare and forecast the data, turning it into information from which strategic decisions are made.

Executives want to know how satisfied their customers are with their company's products. Analytics software enables managers to better understand their customers' shopping patterns, spending behaviors, and product preferences. Many organizations use analytical tools to enhance targeting efforts by segmenting their customer base according to product preferences and potential market value. The cost of automating their business processes to make them compliant with the operational requirements of an analytics system may be prohibitive when considered from the perspective of the benefits that might reasonably be expected to follow. Company executives who continually find themselves unable to respond quickly to changes in the marketplace because their data are outdated or not easily interpretable can benefit from analytics software. Implementing analytical tools contains the following several-step course of preparation before undertaking system deployment:

a) **Process Designing** – understanding where the data currently reside, who is responsible for collecting the information, and how it can be accessed across departments. When, they discover what kind of systems integration is required, they can develop an infrastructure able to be accessed throughout the enterprise.
b) **Data Verifying** – data sources need to be cleaned, standardized, and integrated to ensure that the information will be able to be manipulated as intended by the analytics software. Here, executives should establish who owns the data, whether the data have all been verified, and who will take responsibility for making sure that information is clean.
c) **Implementing the Software** – a key implementation step is working with the analytics software itself to determine how it will interface with other systems,

what data will feed into it, and how individual screens will be designed. Companies typically go through several testing cycles to ensure that the software is working as intended.

d) **Measure the results** – an important element in supporting and maintaining the effectiveness of an analytics software implementation is to measure the results. While the choice of appropriate metrics will vary from organization to organization, one principal reason for deploying analytics remains constant – to increase the effectiveness of customer interactions. The ability to deliver the right information to the right people at the right time helps the world's best companies stay competitive in today's fast changing economy. The primary benefit of any analytics system is its ability to distill vast quantities of data into usable information. As a result, it is clear that a credible learning analytics system must itself be scalable to an enterprise level, limitations on the amount of data that may be used fundamentally limit the value of the system.

3. Business Analytics System Overview

Business Analytics focuses on effective use of data and information to drive positive business actions. The body of knowledge for this area includes both business and technical topics, including concepts of performance management, data visualization, and deployment and use of technology solutions such as OLAP, analytic applications, and data mining. Data mining techniques are used for analyzing and discovering actionable insights from data [Berry 1997]. The use of database systems in supporting applications that employ query based report generation continues to be the main traditional use of this technology. However, the size and volume of data being managed raises new issues. Can we utilize methods wherein the data can help businesses achieve competitive advantage, can the data be used to model underlying business processes, and can we gain insights from the data to help improve business processes? These are the goals of BI systems, and Data Mining is the analytic method that provides the capabilities to explore, summarize, and model the data. Before applying these methods to data, the data have to be typically organized into history repositories – DWs. Data warehousing may require integration of multiple sources of data, which may involve dealing with multiple formats, multiple database systems, distributed databases, cleaning the data, and creating unified logical view of the underlying non-homogeneous data. Data mining analytics try to go beyond OLAP by providing abilities for discovering insights that are computer driven and not end-user driven. Data size is increasing at a rate far exceeding any rates that end-users can cope with. Providing solutions when end-users cannot reasonably supply all possible aggregates to precompute, or when it is not possible to express an insight as a pre-computed aggregate, is the goal of data mining analytics. Analyzing the true business value of websites is becoming increasingly important, but it has also become more complex.

Business Internet Analytics (BIA) is the application of the data mining, data warehousing, and OLAP process to discovering the patterns and rules in the vast data sources.

4. Business Analytics Features

Following are elaborations to the main business analytics features:

DW – a DW is an implementation of an informational database used to store data sourced from an operational database of record. The DW imports a large amount of data collected from several different data sources. It collects day-to-day operational data about users who visit a site: user profile data, transaction data, click-history data, product data, and advertising data. Then manages the data in the database for the purpose of business analytics. DW properties can be configured to specify how data are to be imported.

Business Desk – Business Desk allows analysis of the data collected from websites. The reports and data segments included in the business desk can help an organization make informed decisions. Analysis modules can be used to analyze data about user activity on a website. Specific information about registered users and their behavior during visits to the site can be analyzed, including how much time they spent on web pages and links. The reports can be exported to the targeting system, e.g., a report can be run to find out the number of times each page on the site has been hit or to view the products users have purchased. Reports can be run very quickly because the data required for many of the reports is stored in an OLAP database.

Prediction Models – a model-based algorithm uses a collection of decision trees for making predictions. It is both more accurate and significantly faster than traditional data-based models. A Prediction Model summarizes relationships in the data in the form of rules. Such models make predictions based on the previous activity on the site; consequently, they usually result in more sales.

Segmentation Models – a segment model divides users who tend to have similar characteristics or behavior into segments, which give managers an understanding of the users who visit the site, and can be used for making future decisions. The Segment Viewer module in Business Desk is used to analyze Segment Models.

Business Intelligence (BI) – BI is knowledge gained about a business using various hardware/software technologies. BI enables businesses to turn data into information.

Business Internet Analytics (BIA) – Business Internet Analytics aimed at helping customers gain insight by analyzing customer web data on large, and e-commerce sites [Commerce Server 2002]. BIA allows business users and IT professionals to analyze customer web data and answer complex questions about marketing, advertising, etc.

Data Mining – a technique using software tools geared for the manager who is looking for particular patterns or trends in DW data. Data mining is the process of combing vast amounts of data to produce data content relationships. It can be used to predict future trends and behaviors, allowing businesses to make proactive, knowledge-driven decisions.

5. The Essence for an OLAP

Most companies rely on some form of data warehouse whose vast store of information is organized and interpreted through the application of analytical programs [Sobh 2002]. However, these systems typically are not integrated. Retrieving useful data from them is thus difficult and time consuming. Managers responsible for driving business performance need quick, easy access to real-time sales, marketing, and consumer information to be able to measure the effectiveness of their strategies and make timely changes accordingly. Analytics software, which delivers data from multiple sources and systems in real time, can help executives leverage mass quantities of information in support of rapid and sound business decisions. The software efficiently integrates fragmented data from disparate systems, including front- and back-office, legacy, and Web-based applications.

The proposed solution architecture is comprised of three stages:

a) **Maintenance**

Which were used to integrate and prepare data from various data sources before loading it to the DW – here, all the data were stored in the required format and structure, which was then referenced for historical and up-to-date reporting purposes. Hence corporate memory stores large amounts of data, information, and knowledge from different sources of an enterprise. As transactions occur on a day-to-day basis in learning systems, all or part of the transactions are sent to the learning data warehouse for consolidation and subsequent analysis. The data sent to the warehouse should contain attributes for the various dimensions. In many cases, the dimension attributes will be part of the original transaction. In other cases, the computer programs that are sending information to the warehouse will need to collect and format the dimension attributes prior to sending the transactions along. When the data arrive in the warehouse, they are organized according to the model so that the transactions can be aggregated into measures and accessed across the various dimensions. The DW typically only provide crude tools for selecting and displaying statistics. It is the job of the analytic software to provide the presentation interface and the algorithms to aggregate, compare and forecast the data, turning it into information from which strategic decisions are made.

The system architecture and dataflow of our system is shown in figure 1 below. In terms of our architecture we had done a complete conceptual study of this architecture, however in terms of implementation we are concentrating on the OLAP data mart of the system.

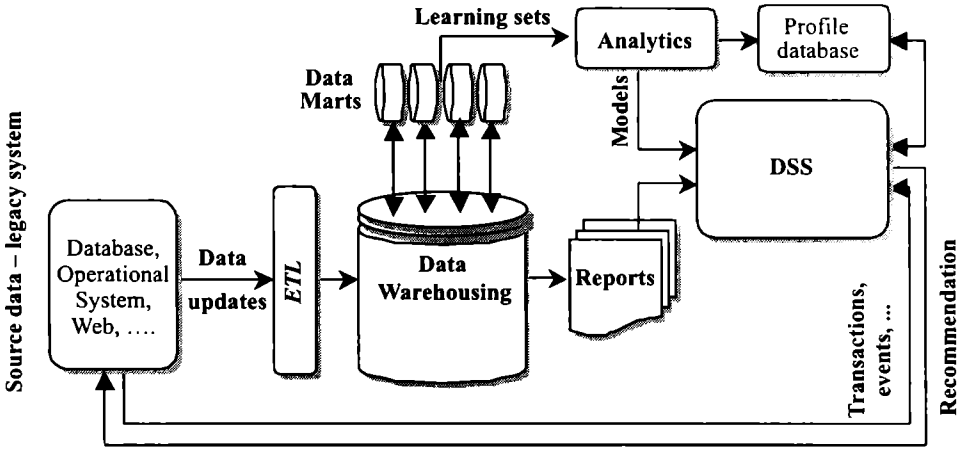

Figure 1. System architecture of our analytics system

b) **Off line analysis**

The best practice strategy, therefore, is to combine with application specific OLAP data marts architecture, where multiple data marts are fed from a central data warehouse. This architecture enables IT organizations to build, maintain and deploy DWs efficiently, and also to meet business users' requirements for analysis and rapid response time. Using a best practice data warehousing/OLAP architecture approach to design a scalable high performance analytics system to be used in any company as a business intelligence solution. Management, sales teams, etc. will use this system to analyze and make key decisions in a timely manner [Matouk 2005]. System functionalities include, but not limited to performance reporting, sales forecasting, product line and customer profitability, sales analysis, market analysis, what-if analysis and manufacturing mix analysis on different levels of aggregation, from different dimensions, and using different graphical types of diagrams. These reporting facilities can be exploited interactively using OLAP technology. Through OLAP the data analyst is enabled to formulate queries and to decide on further queries depending on the outcome of his former queries. In this way, the analyst wanders through the DW collecting information, which he presents to the management. Data are stored in the DW. The data analyst interprets parts of the data, which are represented in a way more adequate for human users. The process of interpreting data needs some knowledge and if the yielded information leads to decisions or actions performed by the management this information becomes knowledge [Karagiannis 2004].

c) **Deployment**

Analytics software offers a customized screen, or dashboard, to help managers gain access to key performance metrics from a variety of applications, including sales and marketing, field service, customer service, and supply chain planning

[Gayeski 2005]. The dashboard, which is usually OLAP technologies, allows the aggregation and summarization of large quantities of data. It aids in the discovery of meaningful data very quickly, and thus helps in making key decisions based on facts This system includes data warehouses, OLAP, analytical reports and the data mining component. users first see data from an aggregate level, then drill down to more granular levels as needed. View details behind the numbers, answering the "Why is this?" question behind the measurements.

6. Conclusion

Data warehousing is a best-in class approach to leverage corporate information. However meaningful business analysis requires a multidimensional view of data that has proven inefficient and cumbersome to express using relational databases. Our analytics solution combines application-neutral DWs with application specific OLAP data marts to combine needs of both IT professionals and various kinds of end users throughout the enterprise. They will be able to extend their decision-support systems by moving beyond a historical focus to proactively chart the future direction of the business. Many organizations already have DW initiatives under way. Unfortunately, integration of learning systems data is not always near the top of the prioritization list. But it should be. Today, data quality is becoming increasingly critical, in the business world, as more and more business decisions are taken based on information provided by databases.

BI is a broad category of applications and technologies for gathering, storing, analyzing, and providing access to data to help enterprise users make better business decisions. Ultimate aim in implementing analytics software is to gain quick access to real time data in order to make better informed business decisions more rapidly. Being able to access and analyze pertinent information on demand, they are likely to be more responsive to market trends, to have deeper insight into customer behavior, to reduce marketing costs, and to gain competitive advantage through better customer service and improved market targeting.

Data warehousing solutions are often expensive and risky because: access and understanding legacy systems require specialist skill sets, every growing volumes of data and demand for real time data feeds, complexity of changing business requirements, source data systems, reporting hierarchies, inability to deliver access to data warehouses for large communities of users, and slow or poor adoption due to little confidence in the data. Several new analytic functions provide basic analytic capabilities within Transact-SQL. These functions will be useful in DWs that allow user queries into the relational database rather than exclusively through Analysis Services. Also, these complex calculations are commonly used during data staging to develop valuable data attributes.

References

Berry M., Litoff G., *Data Mining Techniques for Marketing, Sales and Customer Support*, Copyright by John Wiley & Sons, 1997.

Commerce Server 2002 Capacity and Performance of Business Internet Analytics, [in:] http://www.microsoft.com, Microsoft Corporation, July 1, 2002.

Gayeski D.M., *Managing Learning and Communication Systems as Business Assets*, Prentice-Hall, 2005.

Gonzales M.L., *IBM Data Warehousing – with IBM Business Intelligence Tools*, Wiley Publishing, INC., Indiana 2003.

Inmon W.H., *Building the Data Warehouse*, John Wiley& Sons 2002.

Karagiannis D., Reimer U., *Practical Aspects of Knowledge Management*, 5th International Conference, PAKM 2004, December 2-3,Vienna 2004.

Matouk K., *Podsystem Business Intelligence*, [in:] *Architektura zintegrowanego systemu informatycznego zarządzania*, red. A. Bytniewski, AE, Wrocław 2005.

Sobh T., Valentin H., Koswatta O., *Business Intelligence Analytics System: An OLAP Approach*, ENGR 300, April, 2002.

BIZNESOWY SCENARIUSZ WYKORZYSTYWANIA NARZĘDZI ANALITYCZNYCH W HURTOWNI DANYCH

Streszczenie

Autorzy przedstawiają biznesowy scenariusz wykorzystania najbardziej znanych narzędzi analitycznych w hurtowni danych. Efektywny system analizy wiedzy, zdaniem autorów, musi się opierać na hurtowni danych. Artykuł składa się z dwóch części, w pierwszej został zaprezentowany biznesowy system analityczny (BAS) i jego współpraca z analitycznym systemem OLAP. W drugiej natomiast zademonstrowano najważniejsze zalety, które można osiągnąć przy użyciu tych technologii.

Dr inż. Kamal Matouk jest adiunktem katedry Informatyki Ekonomicznej Akademii Ekonomicznej we Wrocławiu, e-mail: kamal.matuk@ae.wroc.pl.

Mgr Moh'd Al-Sqour jest doktorantem katedry Rachunku Kosztów i Rachunkowości Zarządczej Akademii Ekonomicznej we Wrocławiu, e-mail: mohsqour@wp.pl.