

Data Warehousing (DW) - Models and Business Application

Aida HABUL

Prof. Dr., Sarajevo University,
Faculty of Economics
Sarajevo, Bosnia and Herzegovina
aida.habul@efsa.unsa.ba

Merdžana OBRALIĆ

Res. Assist., International Burch University,
Faculty of Economics
Sarajevo, Bosnia and Herzegovina
mobralic@ibu.edu.ba

Abstract: Fighting with competition requires innovative ways of achieving the advantages in the market. Creating long-term policy of doing business with the possibility of adjustment to unpredictable phenomena requires an enormous amount of quality information about the condition of the company, market, trends of state policies, international trends, etc. This is a huge amount of different data. In the DW this information is extracted, transformed, at the same time detailed and aggregated, processed in a format that suits the user and is available in real time.

The Data Warehouse is a unique picture of business reality and ensures the comprehensiveness of the whole business system; coverage of external and internal data is the basis for defining the business strategy. The DW promotes the business of the company by enriching the business processes and their participants with information needed for making business decisions. It is forced to accurately define and describe the business processes that need to be rejected, imported or innovated. The DW contains rapid, accurate, aggregated, visually accessible information that contains a time dimension, which represents an important managerial resource.

Introduction

In the market there is growing competition that day by day globalizes. At the same time, buyers are more and more selective. All this is forcing manufacturers to constantly check their competitiveness in the market and to find a successful business strategy. Strategic advantage over competitors is sought by achieving the proper planning of market, product innovation, and proper relationships with customers and clients. It is necessary to have timely and relevant information on all aspects of the business and market situations.

Data Warehousing is an important concept of effective decision support systems that are intensively developing in recent years. With this concept and the methods is wanted to achieve "intelligent" business of company in complex market conditions. For this already there is the name "business intelligence".

The company, which analyzes the behavior of their customers, behaves "intelligently". For example, it has all the necessary customer data and captures every activity which is connected to that customer in the database. With the analysis of this data, for example; by analyzing what and when the customer purchased, how many complaints and for which products he had, all these analysis make the customer profile, or can be put into some category. With the analyses of their own business data and external factors, we get the information used in making optimal business decisions, and at the end make profit and ensure the further existence.

Data Warehousing

Data Warehousing is a set of data organization on which a decision support system is based. A small data warehouse that contains the data of only one area is called a regional data warehouse (eng. Data March). Data Warehouse is defined as a set of information organized so can be analyzed, extracted, merged and otherwise be used to understand their essence.

Data Warehousing is also a database, provided that the information in it is organized in a special way. A Data Warehouse has a large amount of data that is organized into small logical units called March Data. Data Warehouse systems use abnormal (relaxed) data. Executing a query over the data organized this way is much faster. This kind of work is called OLAP (On Line Analytical Processing). Data that is embedded in the data warehouse is usually accessed through the March data. March Data usually represent a subset of logically related data from the data warehouse, which refers to a specific area.

Business Intelligence is taken from data generated during the everyday business of firms. Because of the perimeter, this data is archived each year, deleted and only the data of the next year remains in the database. In addition, the data in the database must be updated so that old data is replaced with new data (e.g., old price, the old quantity in warehouses or in stores they are being replaced by new ones), and then the trail of the old data is erased in the database. For operational management of the old values businesses, for example, quantity in warehouse are not important, but they are important for the business intelligence that tracks the time sequence of events of individual business events. As a conclusion we can say that transactional bases are no longer enough, there is a need for data warehouse.

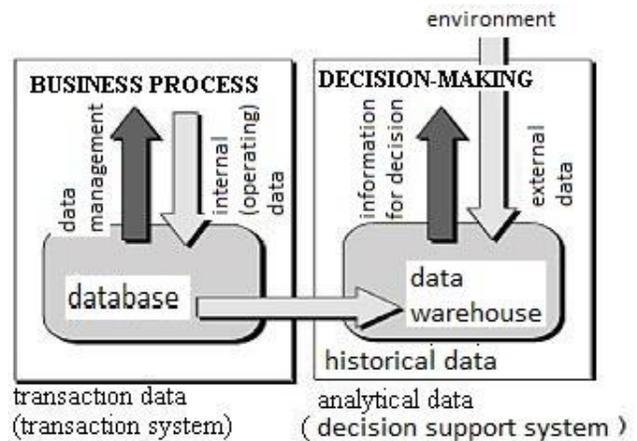
Table 1. view of database and data warehouse

Features of Database

The data in related databases is organized into two-dimensional tables that are mutually connected. The database contains information for each business period. After this period the data is archived and deleted from the current database. The data is constantly being updated and the old values cannot be seen.

What is differet beteen a warehouse and a database?

A warehouse does not accumulate all the data from the database in recent years. Only the data from the database that is thought to be important for the analysis is archived in the warehouse. The data in the warehouse is organized in a multidimensional cube, it is permanent and is related to the time of emergence. The aim of the data warehouse is not operational business, it is to create a richer source of information for various short and long term analysis and for forecasting.



**P r o p e r t i e s o f D a t a
W a r e h o u s e**

The properties of the Data Warehouse are integration of Data Warehouse systems, the orientation of the Data Warehouse system threads, the time dependence of the Data Warehouse system and the permanence of the Data Warehouse system.

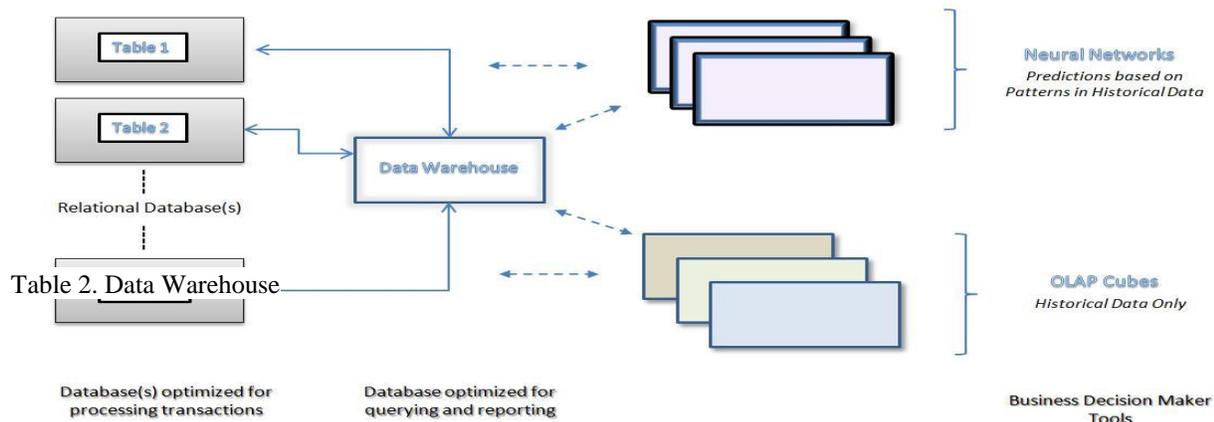
Orientation of DW System on Topics

The data is categorized and organized according to topics of business and not as a functional unit. Examples of possible themes are Sales-information about products, customers, spatial and organizational structure; Marketing-data about markets, products, customers, technologies, plans; Production-data about products, customers, technologies, plans; Financial operations; and Transportation.

Time Dependence of DW System

The data is organized by a series "cross of state" of operational data; each cross refers to a time interval.

How appear Data Warehouse ?



Basic Models of Data Warehouse

When creating a data warehouse today in practice we face three basic models or basic architecture of data warehouse:

(1) two-layer architecture with a common data warehouse (original data + data warehouse); this model is characterizing a single common centralized data warehouse. Data are received from various sources inside of the organization and external data sources available via the Internet or other means. Characteristic of two-layer architecture is that it serves a large number of organizational units of company as individual users. Such warehouses are large scale and very complex, and they are usually stored vast amounts of data. And the scheme of data that are used for storing data should support a wide range application request. From the above it is evident that the costs of maintaining such architecture are high and assume a substantial commitment and time of a certain number and profiles of experts.

(2) two-layer architecture with multiple independent local data warehouse (original data + data march); Characteristics of this data warehouse architecture is the existence of large number of independent local data warehouse intended for support individual applications in the organizational units of companies. The result of this architecture is the large number of systems which each of them submitted their data from various transactional databases. The advantage of the above model of data warehouse is simpler construction and easier to use. But this model also has disadvantages such as: (1) Difficulties in communication among organizational units of companies. This model is not suitable for companies whose business requires the support of applications and projects that assume mutual communication and cooperation of a larger number of organizational units of companies. (2) Increasing the number of mutually independent data warehouse, slow growth and load themselves transactional systems. (3) Data Marches are designed to support only one application and subsequent addition of new applications in a specific warehouse represent difficulties. (4) Limited scalability of platform. And, (5) Uvid into actual state of information at the company level became difficult.

(3) three-layer architecture with a common data warehouse and several affiliated local warehouses (original data + data warehouse + data march);

This model consists of large number of local data warehouse and a common data warehouse that is located between the data warehouse and various data sources within and outside the company. Data warehouses rely on central data warehouse that delivers information in a form that provides uniform access to all segments of company business. Compared to the previous two models the advantages of three-layer architecture are higher accuracy of the information freely from which sources are affected, communication among organizational units is easier, the burden on IT specialists is reduced, stability and scalability of data warehousing platform is increased, and at the end, this architecture is offering possibility of using foreign applications allowing connection of all entities in the chain of value.

This three-layer warehouse architecture, that is known as the most used one has it owns advantages and disadvantages. The advantages are: (1) The work load of IT specialists is reduced (because the data is already coordinated in the warehouse, so they are just taken in the warehouse); (2) More opportunities to add new application; (3) Improves the accuracy and precision of information (because everything is derived from one warehouse); (4) Less difficulty in communication and cooperation between organizational units and parts that has to

be performed together; and (5) The possibility of the use of applications that goes beyond the framework of operations (connecting with buyers and suppliers in the value chain). The disadvantages are: Relatively high cost of establishing a common warehouse at corporate level (to eliminate this disadvantage we have to make a gradual incremental construction of warehouses: first we have to create the first warehouse, then the mini-warehouse, than gradually add new warehouses and activate new applications, and finally comes the independence of the "real" warehouse).

Implementation of Data Warehouse

There are many ways to implement the Data Warehouse in a company. The implementation can be seen from several aspects:

(1) *An area that covers the data warehouse; DW* can contain all the information of one company and for a period of several years. Also, DW can carry a personal character for a specific manager and can contain information from a short period of time (one year).

(2) *Volume of data redundancy*; there are three levels of data redundancy that companies should consider when designing their own DW. (1) Virtual or "Point to Point" DW; (2) Central DW; and (3) Distributed DW.

(3) *Types of final users*; when we look at the DW from the final user point, we can establish three categories of users: (1) CEOs and managers; (2) "Powerful" users (business and financial analysts, engineers, etc.); and (3) Users for support (operational and administrative staff)

Levels of Analysis of Data in Warehouse

1. *Generate static reports* - If we only need a report in tabular or graphical format on the existing data from the data warehouse, report generators are sufficient tools (e.g. Microsoft Crystal Reports or Oracle Reports), which simply display data that is filtered, sorted or summarized by some criteria from the database.

2. *OLAP (on-line processing analytical)* - More complex, analytical processing of data by different dimensions is done at the relational base on OLAP technology (e.g.: if you want to get information about Mazda 323 cars for sale in Osijek Friday afternoon). To view OLAP we have to use the most frequent three-dimensional cubes.

3. *Data Mining* - The most complex part of data processing, means sophisticated Métis for searching hidden laws in data.

Selection of DW Strategies

Before the DW system is developed, it is necessary to choose a strategy that is as far as possible a guarantee that the ultimate solution will meet the specific demands at the time.

Who are the final users? Which areas should be covered? What kind of information should the DW provide? These are just some of the issues raised in the selection of implementation strategies of a DW in a company. We should not ignore the fact that man is the main factor in the decision making process.

Business Application of Data Warehouse

Databases can be seen as the area where the company partners (customers, clients, service users) leave clues about their business relationship with the company. Every transaction, purchase, invitation addressed over the call center, seeking bids for certain set of items is being recorded in database.

Company that are using Data Warehouses

1) Volkswagen AG - uses Business Object tools for monitoring all data - from finance, manufacturing, development, research, sales, marketing and purchasing. Users at all levels can access the statements in BI tools.

2) MasterCard International – uses the BI tools for monitoring data in the area of advertising and authorization and fraud detection.

3) Magma uses OLAP and BI tools to improve the reporting system on the sale of goods.

Example of Applying DW in Textile Trades

A typical retail shop sells textiles to a large number of customers whose needs are significantly different. Without adequate resources for the analysis of numerous data about clients, the shop would be doomed to failure. The biggest benefits for a business are functions, the contact point through which insight into the characteristics and behavior of clients is created.

The development of information technology has enabled us to be informed about clients and also to fill our database with useful information.

The value of a customer is very important data. During their entire life cycle, the value of the customers is always changing. All clients are not equally profitable, so the profitability of customers also changes over time. These are the reasons why it is important to identify the clients whose value during their entire life cycle is great and to establish close relationships with them. Our database should also contain answers to important questions about customers, such as customer loyalty, cross selling, determining target prices, but also about suppliers, the efficiency of suppliers, stock control, product flow, forecasting demand. The emphasis on the importance of logistics leads to specialization in providing logistics services exclusively.

Textile makers are faced with the traditional approach of material offers which a user can feel and with the new generation of customers that manage in the virtual world almost as much as in the real.

A well-made database enables the rapid development and progress in every sense. Also, the path from the manufacturer to the final consumers is reduced to the minimum.

However, the possibilities of keeping the customer in this market are far greater than in others. Today we are witnessing estimates that the market of direct selling as it was known in years past will completely disappear, work on the DW and the content that is offered to customers with quality procurement and distribution channels can ensure the stability of textile makers in new times. Therefore we should not neglect the potential of the Internet, but we should consider target regional environment.

Decisions taken should be closely connected with the database. The database should be well designed, created and filled with good information. That is the only way to guarantee life and successful business.

Conclusion

Fighting competition requires innovative ways of achieving the advantage in the market. Creating a long-term policy of doing business with the possibility of adjustment to unpredictable phenomena requires an enormous amount of quality information on the condition of the company, market, trends of state policies, international trends, etc. This is a huge amount of different data. In a DW this information is extracted, transformed, at the same time detailed and aggregated, processed in format that suits the user and available in real time.

The Data Warehouse is a unique picture of business reality and ensures the comprehensiveness of the whole business system; coverage of external and internal data is the basis for defining the business strategy. The DW promotes the business of the company by enriching the business processes and their participants with information needed for making business decisions. It is forced to accurately define and describe the business processes that need to be rejected, imported or innovated. The DW contains rapid, accurate, aggregated, visually accessible information that contains a time dimension, which represents an important managerial resource.

The DW is an information technology that simply imposes cooperative team work of professional information experts and economists, and links the technical and business worlds. Perhaps this will lead to better cooperation between designers and users of information systems and thus to greater success in the project of building company information systems.

References

Bojan C., Poslovna inteligencija, Beograd, 2006.

Hanic H., Sistem informacija za upravljanje marketingom, Beogradska banarska akademija, Beograd, 2008.

<http://www.infotrend.hr/clanak/2009/7/temelj-procesa-odlucivanja,41,810.html>

John P., Model-Driven Data Warehouse, Integrate Burlingame, CA, 2003

Kellet A., Integrated Business Intelligence, Butler Group, April 2003

Mladen V., Skladiste podataka.

Mrsic L., Poslovna inteligencija kroz poslovne slucajeve, Zagreb, 2003

Mrsic L., Primjena metoda rudarenja podataka u trgovini tekstilnim i srodnim proizvodima, Magistarski rad, Zagreb, 2004.

www.skladistenje.com

www.znanje.org