Trends in Data Processing for Decision Support

Mladen Varga
Graduate School of Economics & Business – Zagreb
Trg J. F. Kennedya 6, Zagreb, Croatia
mladen.varga@efzg.hr

Abstract - The quality of decision-making process is affected by the quality of information used in the process. Trends in data processing for decision support show that business users need analytical applications which incorporate a variety of data analysis techniques incorporating task-specific knowledge, i.e. that are business task-oriented. Such business analytics applications must be vertically connected across the organization, their information is desired to be in a visually acceptable form, even through wireless devices. The analysis functions have to be integrated into the information system, and integrated into the action and result measurements.

I. INTRODUCTION

Every business system, such as an enterprise, strives to establish its information system to provide the information/data necessary for the business system management. Management is the activity or skill of directing and controlling the work of an organization, and in essence it is done by decision-making. To make optimal decisions it is necessary to have complete, reliable and timely information. The paper discusses the trends in data processing: "How to get quality information for decision support?"

It is organized as follows. Business decisions are considered in Section II, software tools used in decision support are discussed in Section III, and the way and trends of data usage are described in Sections IV and V. The information used in decision-making process is settled in the information system which is described in Section VI.

II. BUSINESS DECISIONS

Decision-making is in the base of management. Many authors consider decision-making to be the essence of management, since management is realized only through the decision-making process [1]. Moreover, decision-makers usually accept satisfactory solutions and do not insist on the best ones, since they usually decide in conditions of uncertainty unaware of all possible parameters involved in their selection.

The decision-making process involves almost all employees but the greatest responsibility lies on the organization's management. The importance of the decisions grows with the levels the decisions are made at. Information technology decentralizes decision-making process and makes it more democratic, because it increases the number of participants involved in the process.

Decisions are usually classified as programmed and nonprogrammed. The programmed decisions are used for solving routine problems in repeating situations, in which the decision-making procedure is known. A programmed system, implemented in the transactional information subsystem, can "make" decisions instead of people. For example, granting simple loans can be programmed in such a way that a computer programme grants the loan according to the credit seeker's data. Nonprogrammed decisions are used to solving nonroutine problems in unrepeatable situations in which the decision-making procedure or the decision-making model are unknown. Decision makers take full responsibility for their decisions, but the decision support system may provide relevant data and can support the analysis of this data. The decision support system has to be flexible to function in varied decision conditions and easy to use. Data warehousing systems fulfil these tasks by storing its multidimensional data in the data warehouse and by facilitating analytical processing and data mining.

Regarding their importance, decisions are ranked as operational, tactical and strategic (Fig. 1). Tactical decisions are typical within specific departmental areas and the data needed is usually specific to that department. Some examples of tactical decisions are: "What is the right product mix to offer in a specific market?" and "Is there enough inventory on hand to fulfil the order?" Strategic decisions are usually made by upper management within the organization. They occur less frequently than tactical decisions and can impact many departments and functional areas within the business. Strong data analysis is recommended for strategic decision-making. The analysis often leads to additional questions and additional requests for data. The examples of strategic decisions are: "What distribution channel is suitable for global market?" and "Is it appropriate time to launch a new product on the market?"

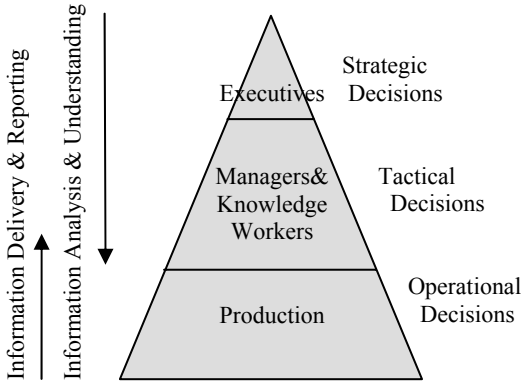

Figure 1. Strategic, tactical and operational decisions

III. SOFTWARE TOOLS IN DECISION SUPPORT

The quality of the business decision-making process is affected by the quality of the information necessary in the process. Better information gives better opportunities to choose the best decision option. Critical information about the state of the business affecting decision-making must be distributed quickly, efficiently, and appropriately to those people and departments that can affect the company's adaptability.

For many years business decision-makers have relied on predefined reports taken from transactional systems. The evolvement of information technology has increased the data available from transactional systems. The reporting backlog has shown results: many reports pulling data from transactional systems rarely provide the decision-makers with the necessary information. Business decision-makers have constantly demanded more information on additional reports. The result was that decision-makers often had to wait for reports too long after they were not provided with the appropriate information. The vicious cycle of demanding new reports took place. In those days some 90% decision-makers relied on reports, less than 10% very technically skilled users were able to work applying analysis tools.

With the arrival of the personal computer technology, individual users, including the entire departments, took over control of their data by using spreadsheets as the central repository of their data. This was unfavourable since the data were usually loosely coupled and integrated across organization, but it also provided an opportunity to a share of users (30%) to easily manipulate their information by spreadsheets. Even today, as massive amounts of information continue to grow, a large portion of data is still analyzed in spreadsheets. Only a small share of users works with data analysis tools.

The majority of decision-makers are still unsatisfied with the information they receive to support their decision-making process. Most reports and data analysis tools fail to comprehend what the data, such as trends, causes of problems, etc., actually means. So the users continue to ask for new reports aggravating the reporting backlog problem.

There is a misbalance of the information delivery model by reporting and data analysis tolls and needs of the decision-making process. The mentioned information delivery model usually cannot answer such questions as:

- What are the most profitable customers? Why are they profitable? How to raise the profitability of the bad customers?
- What are the best sales channels for the new product?
- Where are manufacturing bottlenecks? How can be the bottlenecks efficiently eliminated?

Nowadays we are witnessing the growth of the business analytics software market. Business analytics allows organizations to go beyond the traditional business intelligence applications by providing an integrated, enterprise-wide view of information and a higher level of predictive insights and optimization. The key challenge is to deliver this knowledge through solutions that are pertinent to specific business processes, enabling decision-makers to achieve greater return on investments and organizational efficiencies. Business analytics software may contain hundreds of pre-built metrics, reports and alerts. For example, marketing analytics may help create precise segmentation strategies using all relevant data, including historical, financial, demographic, and syndicated data to execute highly effective up-sell, cross-sell and retention campaigns. It may also help create compelling customer dialogue through individually targeted, event-based marketing communications based on customer life event changes, purchase behaviour, service history, or predicted response. It may also track and analyze marketing campaign performance and help marketers make timely course corrections based on which offers and messages are driving the highest response rates.

An enterprise's information assets, such as customer data, can be vast. Too often they are thrown away in information silos and never used. The goal of business analytics is to turn these marginalized data resources into information that allows business managers grasp the dynamic state of their business.

Business analytics is moving beyond data warehousing, which a limited number of experts usually use, to include other users, possibly on publish-and-subscribe basis, to distribute business intelligence.

The challenge of business analytics is to capture external data from sources that companies have not considered before. There are many sources of data available - many of them via the Web.

The view of the business also changes. Organizations want a single view of the customer, particularly given the current economic trends. This requires a move away from point solutions to more integrated systems.

The mentioned trend of using software in decision-making support is graphically depicted in Fig. 2.

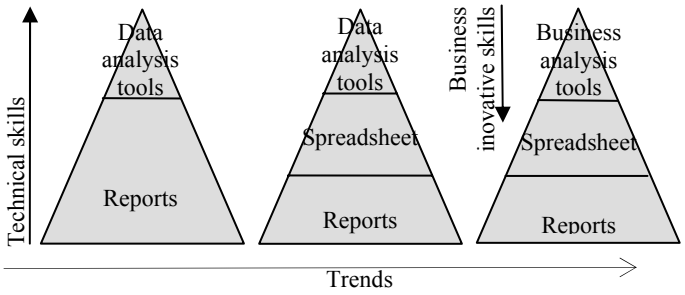

Figure 2. Trends in using software for decision support

IV. DATA USAGE IN DECISION SUPPORT

To find out what is necessary for effective decision-making support tool we need to enter the world of decision-makers. Whoever the business decision-makers are, they make decisions regardless of what tools they have and whether the information they receive is useful or not. Therefore, it is essential to understand the fundamentals of the decision-making process.

A decision-maker uses a variety of different methods to use, i.e. review and analyse data (Fig. 3). They differ by the character of data usage - from simple, i.e. reading, to complex, i.e. predictive, analysis.

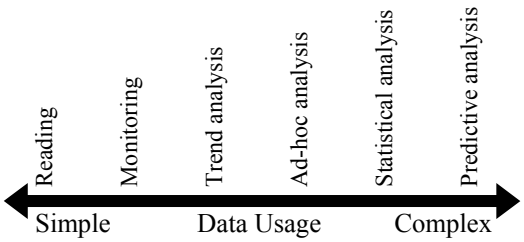

Figure 3. Character of data usage

It is important to understand exactly how decision-makers interact with the information prior to making decisions. Business decision-makers use data in the decision-making process in various ways depending on the type of task they are involved in. Table 1. shows what type of data usage is used to fulfil what type of tasks.

Table 1. Data usage and tasks

Type of data usage	Type of tasks	Examples
Reporting and monitoring	*Based on rules*: Simply looking at a pre-defined report for specific information pertaining to the subject; monitoring information at a specific period of time	The user reads inventory report through products looking for quantity on hand. The user opens daily business reports to understand how the business is progressing.
Analysis (elementary analysis)	*Based on skill*: Delivering more complex information the most often in iterative manner	The user analyses trends - new information seen in context with historical data in order to understand change in time.
Knowledge discovery (sophisticated analysis)	*Based on knowledge*: Analysing data combined with problem modelling (data mining) to understand more deeply information and uncover hidden insights into data	The user carries out affinity analyses, such as basket analysis to understand consumer behaviour.

A. Data Usage Based on Rules

Reporting and monitoring are examples of data usage based on rules. Reporting technology is intended for information distribution. Both predefined and ad-hoc reports are addressed at information delivery, publishing and distribution. This is the reason why reporting tools have menus with menu items that predefine the query and publish the report to the targeted users.

The transactional or operational part of the information system is responsible for executing the transactions of the enterprise's business processes. A part of it is the management reporting subsystem consisting of a set of predefined and standardized reports based on transactional data. The important factor of the quality of reporting subsystem is the degree of data integration across enterprise's functional areas. The successfully integrated enterprise's data is the main factor of the success of integrated information systems (Enterprise Resource Planning – ERP and Enterprise Application Integration – EAI).

Types of decisions (programmed decisions), such as decisions regarding credit approval or inventory on hand, may be incorporated in the transactional subsystem.

Nowadays decision-makers are provided with predefined standardized and parameterized reports or are forced to make ad-hoc query via reporting tool to make necessary reports. However, the complexity of data, the variety in data encoding, calculation and the need to learn more advanced decision support tools often deter decision-makers from using advanced decision support tools. Therefore, the users leave with predefined reports and with no access to deeper information. This is the main cause for decision-makers' frustrations with the level of information support in the decision-making process. Decision-makers receive reports of all shapes, but never feel they have the exact information needed for decision-making. The situation is described as "data rich but knowledge poor".

The type of data containing the information usable in decision-making processes is semi-structured data. Its content consists of various databases or document bases populated with business documentation, business correspondence, business messages, HTML or XML data, etc. Content management is often used to refer to the process of creation, management, distribution, publishing, and discovery of corporate information. Web content management typically focuses on online content targeted at either a corporate website or intranet. Document management assists organisations to manage the creation and flow of business documents.

In the corporate business decision-making process the important role is played by collaborative/groupware systems that allow their users to collaborate with each other.

B. Data Usage Based on Skills

Analytic solutions are examples of data usage based on skills. They are different from reporting solutions as they are intended to facilitate the analysis of information such as inspection, exploration etc., in order to assimilate and understand information. Menu items in analysis tools facilitate information exploration, information patterns, historical content analysis etc., and use advanced data visualization to help users to evaluate the relevance of the information.

The analytic system must help the decision-maker understand the information and the information insight. It assists the user in getting the answer "why" and "how" behind the information given in an analysis report. An example of this kind of analysis is "The sales are lower than before, why? Is it because of a specific product, a specific market, a specific customer or a combination of all of them?"

The part of information system enabling corporate decision-makers to supply information and thus navigate

the complex business environment is the analytical or decision support subsystem. Today it is often referred to as business intelligence.

Dimensional analysis of data, such as On-line Analytical Processing (OLAP) in data warehouses, is a good example of analytics. People intuitively look at the business through dimensions or perspectives. The managerial question "What are the sales data on products, time and markets" implies three business dimensions: product, time and market. The sales attributes are referred to as measures or facts. The facts are mostly numerical, preferably continuously valued and additive. They vary over time. In the statistical database field [4] the dimension corresponds to the category attribute, and the measure corresponds to the summary attribute. The cube model is very useful in presenting the concept of the multidimensional space.

Dimensions are usually organized into hierarchies that specify aggregation level and hence granularity of viewing data. For example, the hierarchy of the product dimension is Product→ProductGroup→ProductType. The dimension hierarchies are important in drill-down and drill-up operations, i.e. in the analytic/synthetic view of the dimensional data.

The multidimensional (analytical) data structure is in many aspects more visual than table structure used in operational (transactional) information subsystem. The reason is that the dimensional structure is capable of showing interesting interrelationships between dimensional attributes, such as product, time and market, which is not seen in classical table structure. These interrelationships among the dimensional attributes may result in valuable business information. For business decision-makers they are the most important because they may uncover hidden business information.

The processing of multidimensional data is characterized by a set of specific functions, such as dimensional rotation, dimension position selection, drill-up and drill-down through dimension hierarchies, etc.

C. Data Usage Based on Knowledge

The best example of data usage based on knowledge is data mining. Data mining [5] is an analytic process designed to explore usually large amounts of business data in search of consistent patterns and/or systematic relationships between variables, i.e. this is an information extraction activity whose goal is to discover hidden facts contained in databases. Data mining is a combination of machine learning, statistical analysis, modelling techniques and database technology.

The goal of data mining is prediction. Predictive data mining is the most common type of data mining. It has the most direct business applications. Typical applications include market segmentation, customer profiling, fraud detection, evaluation of retail promotions, and credit risk analysis.

According to [5] the process of data mining consists of three stages: (1) the initial exploration, (2) model building or pattern identification with validation/verification, and (3) deployment, i.e. the application of the model to new data in order to generate predictions.

The exploration stage involves cleaning data, data transformations, selecting subsets of records and variables Depending on the nature of the analytic problem, this stage usually involves a wide variety of graphical and statistical methods known as exploratory data analysis, with the purpose to identify the relevant variables and the general models that can be used in the next stage.

At the model building and validation stage various models are considered. Test best is the model with the great predictive performance which most accurately explains the variability in question and which produces stable results across samples. This stage may be a very elaborate process using a variety of techniques developed to achieve that goal. Different models may be applied to the same set of data and then compared to choose the best. These techniques are referred to as predictive data mining, statistical learning, etc.

At the deployment stage the best model selected in the previous stage is applied to new data in order to generate predictions or estimate the expected outcome.

Some other methods, such as optimization and simulation, are used to rarely in business decisions-making process. They facilitate finding optimal solutions of various problems. Here are examples of some methods used in supporting the decision-making process. The linear optimization methods, such as linear programming, integer linear programming, transport problem, etc., help finding optimal solution to the problem described by linear functions. Multi-criteria decision-making methods, such as Analytic Hierarchy Process (AHP), are used in selecting alternative solutions to the problem. Simulation methods, such as discrete event simulation and system dynamics, are used in complex systems with a great number of mutually connected elements and time variant behaviour. They solve problems which cannot be modelled with mathematical methods.

Expert systems can solve a narrowly defined set of problems using information and reasoning techniques normally associated with a human expert. Expert systems are intelligent systems capable of using knowledge to solve the problems. They consist of a knowledge base and a reasoning mechanism.

V. TRENDS OF DATA USAGE IN DECISION SUPPORT

Both reporting and analytics (elementary and sophisticated) together provide an organisation with the combination of data and necessary tools to understand information. When these two technologies work together, users can get information delivered to them in a variety of ways. When users need to know more about information, they can launch analysis and determine the factors that are critical to optimal decision-making.

The key issue is to pick a reporting system that satisfies all users' needs and then pick a data/business analysis system so they can work seamlessly together.

A. Business Users

Although business users are experts in their business domains, they are unlikely to be experts in data analysis [6]. Usually business users rely on a data analyst who employs analysis applications to extract information from

data. Business users have to impart their domain knowledge to the analyst, and then wait until the analyst organizes data, analyzes it, and returns the results. Since there are usually open questions regarding the results, several iterations are necessary before business users can act on the results of the analysis.

Analytical applications, which incorporate a variety of data analysis techniques, must therefore provide recommendations to business users of how to best analyze data and present the extracted information for the specific business problems. From the business users' standpoint business analytics must rely on solving specific business problems, i.e. it must incorporate task-specific knowledge, and must not rely exclusively on data analysis techniques. This the reason why this type of analysis is called business analytics.

B. Verticalization

An important trend seen in business data analysis software is verticalization [6], i. e. incorporating specific task knowledge in the software. An illustration of this is a marketing campaign and the analysis of customer data in improving the effectiveness of the campaign. Since business users are not fond of statistical concepts, analytical systems have started to utilize visualization of the result and have headed the results to task-oriented outputs.

C. Availability

The way the extracted information is presented and accessed is of paramount importance to business users. In some industries it is increasingly desired that the business analytic information is available through wireless devices in a visually acceptable form.

D. Analytics integration

Analysis functions are now being incorporated into information systems, instead of being outside the system. Regarding the type of data analysis they are incorporated into the decision support subsystem or even into the transactional part of the information system. Integration of data from multiple data source is very important. Analysis is more effective when data is available from multiple sources. As an example, customer data is more powerful if overlaid with demographic data. Another example is integration of Supply Chain Management and Customer Relation Management allows enterprises to integrate demand and supply chains in optimal manner.

E. Integration of analytics with action and result measurements

Two important questions [6] connected with data analytics are: "How to turn discovered information into action?" and "How to know the effect of each action?" It is necessary that solutions use analytical results as a starting point towards the next steps actions and measurements of the result. Integration between analytics and operation system is the key. For example, analytical application is identified as a cluster of customers who respond to a promotion, and the operation system responds with the distribution of catalogue with optimized promotions.

VI. INFORMATION SYSTEM

Probably the most information used in business decision-making process can be found in the organization's information system. The information system is a complex system that has to cover all informational tasks needed to service operational, management and decision-making activities of the enterprise [2]. It may be decomposed into three information subsystems: transactional or operational subsystem, analytical or decision support subsystem, and collaborative subsystem [3].

Business intelligence and business analytics rely on the information system. Table 2. shows information subsystem's content, structure of the content and the aim of processing the content:

- Transactional (operational) subsystem maintains the content in the table (relational) structure to help the business process execution
- Analytical (informative) subsystem maintains its data in the data warehouse in the dimensional structure to help the decision-making process
- Collaborative subsystem maintains its semi-structured data in the document base to help the collaboration and communication

Table 2. Information subsystems

Information subsystem	Transactional (operational)	Analytical (informative)	Collaborative
Content	Data base	Data warehouse	Document base
Content structure	Table structure	Dimensional structure	Semi structure
Aim of the content processing	Business process execution	Decision-making, management	Collaboration, communication

All kinds of information can be gathered and aggregated in information systems, but if there is no context of it and if business has no advantage to generate a good business strategy, then information is worthless. Information system's data must be integrated vertically and horizontally, users must know the data stored in the information system and the data must be documented. Not only internal data has to be integrated, this must be true for both interior and exterior data.

The classical transactional systems are designed to shorten the business process cycle. Nowadays, business intelligence and business analytics shorten the decision-process cycle and increase the productivity of expert and knowledge workers.

VII. CONCLUSION

Business intelligence and business analytics deal with supplying the information for the decision-making process. Their applications try to put the right information at the right time into the hands of right user. They try to transform data into information; and information into knowledge in order to make the right decision and finally gain the profit.

Trends in business analytics are driven by business value, which is measured as progress towards bridging the gap between the needs of business decision-makers and the usability of analytic tools.

Current trends show that business analytics applications must incorporate task-specific knowledge, vertically connected across the organization. Business analytics information is desired to be in a visually acceptable form, even through wireless devices. Analysis functions have to be integrated into information systems, and integrated with actions and result measurements.

Business intelligence and business analytics apply various data processing methods in order to help the organization to reach intelligent behaviour in turbulent business environment. Unfortunately, this cannot be bought; it has to be built up patiently.

Business intelligence and analytics tend to reduce the number of management levels cancelling middle level managers. Business intelligence designers have a hard task to synchronize the procedures and data between the strategic and the operational management levels, and to exclude the requirements of middle management who want to reproduce the network of unnecessary relationships.

LITERATURE

[1] P. Sikavica, B. Bebek, H. Skoko, D. Tipurić, Poslovno odlučivanje, Informator, Zagreb, 1999.

[2] M. Varga. Motivation Issues in the Framework of Information Systems Architecture, Journal of Information and Organizational Sciences, vol. 27, no. 2, pp. 109-118, 2003.

[3] M. Varga. A Contribution to Information Systems Courses Taxonomy. Proc. of the 15th International Conference on Information and Intelligent Systems, p. 61-68, Varaždin, 2004.

[4] A. Shoshani, OLAP and Statistical Databases: Similarities and Differences. 16th ACK SIGACT-SIGMOD-SIGART Symp. on Principles of Database Systems, p. 185-196, 1997

[5] Data Mining Techniques, Statsoft Inc., http://www.statsoft.com/textbook/stdatmin.html [2005-01-28]

[6] R. Rohavi, N. Rothleder, E. Simoudis, Emerging Trends in Business Analytics. Communications of the ACM, vol. 45, No. 8, p.45-48, 2002.