Emerging Data Management Tools and Their Implications for Decision Support

Sean B. Eom* · Elena Novikova* · Sangjin Yoo**

Abstract

Recently, we have witnessed a host of emerging tools in the management support systems(MSS) area including the data warehouse/multidimensional databases (MDDB), data mining, on-line analytical processing(OLAP), intelligent agents, World Wide Web(WWW) technologies, the Internet, and corporate intranets. These tools are reshaping MSS developments in organizations.

This article reviews a set of emerging data management technologies in the knowledge discovery in databases(KDD) process and analyzes their implications for decision support. Furthermore, today's MSS are equipped with a plethora of AI techniques (artificial neural networks, and genetic algorithms, etc.), fuzzy sets, modeling by example, geographical information system(GIS), logic modeling, and visual interactive modeling(VIM). All these developments suggest that we are shifting the corporate decision making paradigm from information-driven decision making in the 1980s to knowledge-driven decision making in the 1990s.

Keywords: data warehouse; data mining; knowledge discovery in databases; management support systems; multidimensional analysis; on-line analytical processing.

Ⅰ. Introduction

Recently, we have witnessed a host of emerging tools in the management support systems(MSS) area. They are becoming an integral part of a set of recent developments in data management, decision support systems(DSS), and executive information systems (EIS)/executive support systems(ESS). These tools are reshaping DSS developments in organizations. These emerging tools include the dat warehouse/multidimensional databases (MDDB), data mining, on-line analytical processing(OLAP), intelligent agents, World Wide Web(WWW) technologies,

* Southeast Missouri State University

** Keimyung University

the Internet, and corporate intranets, These new tools are a set of inseparable tools that adds new capabilities to DSS, EIS, and ESS.

The data warehouse is a relational/ multidimensional database which is separated from operational databases. It is a subject oriented, integrated, time-variant, and nonvolatile(read only) collection of databases optimized for decision support. MDDB organizes data as an n-dimensional cube so that users deal with multidimensional data views such as product, region, sales, time, etc. with a faster query response time. Data mining, as a subactivity in the Knowledge Discovery in databases, refers to discovering meaningful information/ patterns/ trends about a business from the data warehouse that queries and reports don't reveal effectively using various techniques(Gray 1996).

Intelligent agents(known also as intelligent interfaces, adaptive interfaces) research is an emerging interdisciplinary research area involving researchers from such fields as expert system, decision support systems, cognitive science. psychology, database, etc. According to Riecken(1994), the primary purpose of agent research is to "develop software systems which engage and help all types of end users" in order to reduce work and information overload, teach, learn and perform tasks for the user.

In the 1992 Franz Edelman DSS prize-winning paper, Angehrn(1993) introduced the conversational framework for decision support. The conversational framework is the basis of a new generation of active and intelligent decision support systems and executive information systems. The active DSS will be equipped with the tools (stimulus agents) that will act as experts, servants, or mentors to decide when and how to provide advice and criticism to the user, while the user formulates and inquires about its problems under the continuous stimulus of electronic agents. This kind of active DSSs promotes use, creativity, exploratory learning, and adaptability. The essence of active decision support activities includes monitoring decision making processes and stimulating creative ideas through carrying out insightful conversations with decision makers.

World Wide Web-based DSS is another emerging topic in the DSS area. The World Wide Web is increasingly being used as the client/server platform of many business organizations due to its network and platform-independence and very low software/ installation/

maintenance costs. The web-based solutions are low cost vehicles for easily accessing, analyzing, and distributing timely business information from corporate databases through OLAP. The Internet and corporate intranets opened a wide possibility of building decision support systems to deal with problems of global natures. As we enter the age of the global village where geographical and temporal boundaries are shrinking rapidly, global decision support systems are emerging as the frontiers in management information systems area (Eom 1996a). Many multimillion dollar research projects are being launched to develop, implement, and evaluate WWW-based decision support systems in the health care area (Anonymous 1997).

Due to the digital revolution, more and more data are easily and automatically captured and accumulated through the use of source data automation devices such as point-of-sale systems, optical bar-code-scanners in supermarkets, and optical charter recognition devises in banks, The massive amount of data now can be stored in storage devices with manageable costs. In addition, developments of parallel processing technology have also contributed to the emergence of the data warehouse, data mining, on-line analytical processing, etc.

This article examines the two critical questions. How do we organize the massive amount of data for effective decision making and how do we automatically process them in a meaningful way to create knowledge? The data warehouse is being implemented by many organizations to answer the first question and knowledge discovery in databases and data mining are emerging topics in the DSS/EIS/ESS areas to answer the second question. The article also analyzes the implications of a set of emerging data management technologies in the knowledge discovery in databases (KDD) process for the researchers and practitioners in the DSS/EIS/ESS areas.

II. The KDD Process

Knowledge discovery in databases (KDD) is defined as " the overall process of discovering useful knowledge from data." (Fayyad, et al, 1996, p. 28). Data mining is a particular step of applying specific algorithms and other tools for extracting patterns/trends/relationships from data. The significance of the KDD process is that this new development attempts to automatically extract knowledge from data in the large databases, either in the data warehouse or elsewhere.

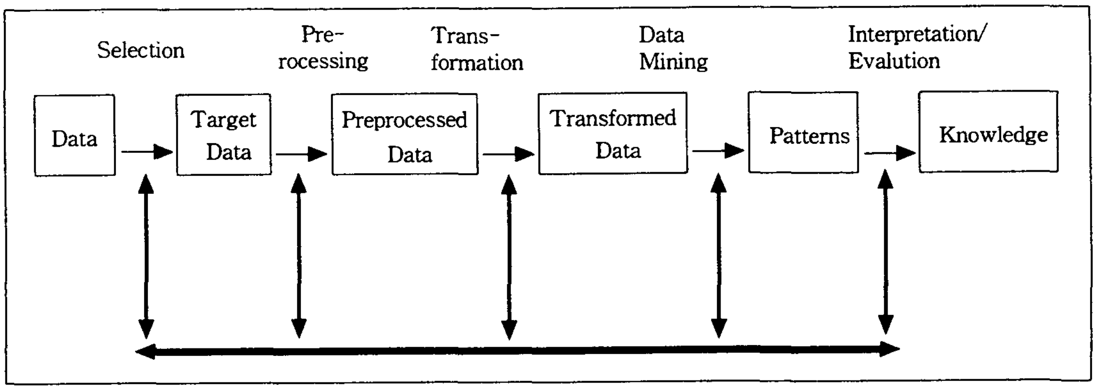

Figure 1. Overview of the steps constituting the KDD process

1. The Data Warehouse

The KDD process does not have to begin with the data warehouse. The KDD process can be carried out on any form of data such as spreadsheets, transaction processing systems files, etc. However, Data stored in the data warehouse are organized to support decision making to include integrated data, detailed and summarized data, and historical data, and historical data, and meta data. Therefore, there is a symbiotic relationship between the data warehouse and data mining (Inmon 1996).

1.1. Definitions of Data Warehousing and the Data Warehouse

A data warehouse is a separate database from an operational database created for management information and decision support. Many practitioners view the data warehouse as " the architectural foundation of decision support systems" (Inmon 1996) or " a total information and DSS architecture" (Chevinsky and Henderson 1997). The main purpose of a data warehouse is to give timely and useful information about the organization, its products or services, and its customers to decision makers to help them effectively manage the organization. In doing so, it is imperative to build a separate data warehouse for effective and efficient decision making which need cleansed, summarized and time-based data. More importantly, an operational database is optimized for transaction processing. For efficient end user query operations, a separate database is usually extracted. A separate data warehouse prevents on-line transaction processing systems from having degraded performance (e.g., slow processing time or reduced response time).

It is estimated that the average first pilot of a full data warehouse may cost $600,000 and that an additional outlay of approximately $1.3 million may be required during the first three years. Therefore, an important question before making the decision to build the warehouse is: Do you need a data warehouse? The answer is "yes" if your organization has: a desire to stay close to customers; information-based products or services; highly competitive markets; too much data stored in many different systems; and a need for more effective " ad-hoc" data querying and reporting. (Chervinsky and Henderson 1997).

Data Warehousing is the process of developing an enterprise wide data warehouse, as defined below.

According to Inmon (1993, 1996), the founding father of data warehousing, data warehouse is "a subject oriented, integrated, nonvolatile, time-variant collection of data organized to support management needs."

(1) Subject oriented - the data is organized by subject or entity of the enterprise such as customer, vendor, supplier, product, etc. A data warehouse organizes all the data into difficult subdirectories such as products, customers, inventories, etc. It is able to receive pieces if information from various sources (such as departments of an organization) and store them in corresponding subdirectories.

(2) Integrated - The data is entered in a consistent format. For example, in a sales application, a product may be measured in lineal meters, whereas the same product may be measured in square meters in the inventory control systems. In the data warehouse, only one of the two must be chosen to represent the units of measuring the same product. Data can be integrated/ cleaned/ conditioned by standardizing data structures, reconciling encoded values, reconstructing keys, etc. in a single, globally acceptable format regardless of multiple data sources. Data are integrated through denormalizing tables, cleansing data of error and redundancies, and adding new fields and keys to reflect end-user needs for sorting, combining, and summarizing data.

(3) Non-volatile - The data in the warehouse cannot be updated. The data

can only be read (accessed), but it cannot be over written (changed). The warehouse data is static(long-living) while operational data is rapidly changing(short-lived).

In addition, Gray(1996) added several additional properties of the data in the data warehouse.

(4) They contain time series data. Consequently, the data do not show current status.

(5) They are summarized. They contain both aggregated and atomic data (relevant to some moment of time) so that they can be used in the decision making process directly and immediately. Data are aggregated into sets, which is why warehouse data is friendly to relational data bases.

(6) They are larger in size. The size of the data warehouse can be very large. Some financial institutions have data warehouses that take up to two terabytes of disk space(Newing 1996 b).

(7) They are not normalized and therefore can be redundant;

(8) They contain metadata (data about data); Metadata is an integral part of data warehouse. This database contains information about when the information was created or updated and who was responsible for that. It also has the information on warehouse data structures, integration logic and process. It shows how data is related on different levels, and it links everything together in the architecture of a data warehouse (Inmon 1993).

1.2. Definitions of Data Mart

A data warehouse can be useful for any organization, but considering costs and benefits, only big companies with large volume of data are using it (Darling 48). Also, many organizations do not have one big warehouse. They rather separate their data into smaller warehouse for smaller number of users and limited content, such smaller warehouse is called data mart(Castelluccio 1996). As users work with data(for example, performing trend analysis or what-if analysis) they create more and more data. A data warehouse is able to save this data

1.3. Data Warehouse Architecture

"An architecture is a set of rules or structures providing a framework for an overall design of a system or a product" (Poe 1996). Data warehouse architecture has several components: data model, warehouse server, application server, middleware layer, data-scrubbing utilities, data-transport utilities , replication engine, meta-data repositories, client analysis tools (Baum 1996).

The simple form of data warehouse architecture consists of two parts (tiers); data warehouse is located on single server tier and analytical and decision-support tools are located on client tier. The more sophisticated data warehouse architecture consists of three parts (tiers): the server tier with data warehouse itself, the middle tier is any type of scheduling program or a multidimensional OLAP engine, and the last tier is the client tier which includes decision support and data presentation tool. (Baum 1996).

An essential element of a dat warehouse is decision support tools. The tools are often called as the decision support engines/decision support systems engines. They are general-purpose front-end query tools to construct, execute, and display the results of ad hoc queries, in addition to generate periodic strategic reports and system-triggered exception reports.

III. Data Mining

After the data warehouse is created (as figure 1 shows), the subsequent steps are selecting a targer dataset(a subset of the data in the warehouse) on which knowledge discovery activity is to be performed, preprocessing the dataset which includes cleaning the target dataset(e.g., removing noise, handling missing data fields, etc.), transforming the preprocessed dataset to reduce the number of variables, and data mining.

1. Definitions of Data Mining as a subactivity of the KDD Process

As defined earlier, knowledge discovery in databases is " the overall process of discovering useful knowledge from data." (Fayyad, et al, 1996, p.28). Data mining is a particular step of applying specific algorithms and other tools in the KDD process for extracting knowledge(patterns/trends/relationships) from data. Data mining is the process of discovering new knowledge from detailed data, using a set of analytical techniques, that support decision making. The real advantage of data mining is that using historical and factual data, it creates more knowledge from examples (historical data). Data mining finds new patterns, trends, relations, classes, and insights that are hidden in historical data and that cannot be recognized by such traditional query tools.

Data that are gathered by automated data entry devices such as bar-code reader/scanner, optical character recognition scanner, manager ink character recognition(MICR) device, smart card reader in point of sale (POS) terminal, etc. creates additional difficulties for human cognitive abilities. Such data is highly complicated. There is a great number of association that human mind is not able to comprehend. Some data are hidden, the relationship between data elements are difficult to be seen and analyzed. Data is highly distributed (human mind is not able to see relationships).

2. Data Mining Tools

The data mining technologies utilize a plethora of tools such as artificial intelligence (expert systems, neural networks, pattern recognition, machine learning, and fuzzy logic), statistical analysis, multidimensional data analysis, visualization including geographical information systems, and database monitoring technologies to discover new information, patterns, and trends from a company's databases. OLAP, also known as multi-dimensional analysis, is a technology that allows manipulation of enterprise aggregate data across many dimensions such as product, time, and location, etc. (Codd, et al 1993). Database monitoring uses various tools such as intelligent agents, machine learning, database triggers, and web crawlers to automatically identify data and events to be monitored the specific thresholds that will trigger an alert, and present the results of monitoring to the user.

Data mining techniques are mostly from artificial intelligence (AI) and statistics. The field of machine learning (a branch of AI) aims at incorporating the ability to learn into software. Researchers in the machine learning area have long pursued to create computer systems/programs that can learn from experience/historical cases. A large number of AI learning, cased-based reasoning and analogical reasoning, genetic algorithm, and neural networks.

The majority of data mining tools are applying machine-learning algorithms and use inductive reasoning. Machine learning algorithms derive rules by examining a large number of sample data in a database. For example, a data mining program processed a large volume of retail data and discovered that 75 percent of the customers that purchases scuba gear take Hawaiian vacations. The data mining program extracts an association: if the customers buy scuba gear, they will most likely take Hawaiian vacations.

Table 1. Comparing the Leading Data Mining Technologies.

Technology	Advantages	limitations
Rule-based analysis	Good for data that is "complete" with data relationships that can be modeled in via if .. then rules or decision trees. Rules are readable.	Large number of rules are difficult to understand. Data may not have strong rules-based relationships.
Neutral networks	Good for data with non-linear relationships. Can work well if data is missing some values.	Inability to explain the found relationships, although some leading-edge tools are attempting to create explanations of the decisions. Requires non-numeric data to be converted to numeric data values.
Fuzzy logic	Can rank results based on closeness to the desired result.	Small number of applications and vendors.
K-nearest-neighbor	Good for discovering clusters; can utilize an entire data source rather than require sampling for training.	Requires a large amount of memory (this technology is also called memory-based reasoning). May be overly sensitive to closely matching records.
Genetic algorithms	Good for forecasting problems involving data with non-linear relationships. Can work well if data is missing some values.	Inability to explain the found relationships, although some leading-edge tools are attempting to create explanations of the decisions. Requires non-numeric data values to be converted to numeric data values.
Advanced visualization	Users control the discovery of relationships via highly sophisticated 3-D visualization.	Requires more user intervention than other methods.
Combination	Will choose the best technology for the problem or can compare results of the different technologies. Users are not required to learn several tools.	Can be complicated to use the tool because of its complexity. May not provide best-in-class techniques for each technology.

(Brooks 1996).

Consequently, the data mining program generates the following rule:

IF Customers buys scuba gear THEN Customer will take Hawaiian vacations with 75 % probability.

It uses techniques such as neutral networks, clustering methods, machine learning algorithms (also known as rule-induction software), genetic algorithms and decision trees like Classification and Regression Trees (CART) and Chi Square Automatic Interaction Detection (CHAID) (Newing 1996).

3. Business Applications of Data Mining

Data mining is a part of a knowledge discovering process, though it is still in its early stage, there are several useful business applications through the identification of the following several hidden functional relationships among data items.

3.1. **Associative Relationships**: Market basket analysis has been a successful application of data mining tools. It involves analyzing relationships between things done together such as grocery items that were purchased at the same time. For example, an analysis of a retail database in Korea may generate the following association between two products.

IF customers purchase whisky, THEN customers will purchase soju with 80% of probability.

An analysis of a retail database in the U.S.A. may generate the following association between two products.

IF Men buy diapers, THEN they will buy beer with 65% of probability. Another example of data mining tools using associative relationship is in the area of fraud detection. Abuse is hard to detect and find; however, data mining applications can provide hypotheses or suggestions that there might be a fraud, but further human investigation is necessary. Insurance companies in the U.S.A. have launched high-tech assault on fraud. The National Insurance Crime Bureau manages a comprehensive data warehouse of about 300 million insurance claims records collected from a worldwide information system network of nearly 1000 organizations (insurance companies, vehicle manufactures, law enforcement agencies, etc.). Proactive fraud detection system called ClaimSmart system of the

Bureau use the KDD process to find patterns of claims activity(Abbort 1996).

3.2. **Sequential Relationships**: Data Mining tools analyze sequential relationships of events over time, e.g., purchases of house and dish washer.

Finding these associative and sequential relationships can support various decisions such as shelf-space allocation, store layout, formulating market segmentation strategy, formulating marketing mix strategy, etc. Data mining becomes a very effective tool in that businesses shift their focus from mass merchandising marketing strategies to more customer-oriented strategies, so companies need more detailed information about their relationships with customers.

3.3. **Clustering Data Items**: Data mining tools also useful to cluster the entire data set into several subsets according to several variables (or attributes) of the dataset. Credit scoring is an example of the application based on clustering data items. Data mining systems can be used to predict who among current customers will pay on time and who will fail to make their payments.

3.4 **Predicting based on time series data**: Data mining can be used to can forecast the future behavior of potential customers based on the behavioral patterns of current customers.

3.5 **Classifying** each record of the database into two groups(or more than two groups) such as buyer or nonbuyer. Machine learning tools can process every field in every record in the database and identify the characteristics between two groups of data. Here is an example how database mining system was able to identify very useful relationships for a bank. The system analyzed the characteristics of all customers and came to the conclusion that customers that have a tendency of overdrafting are excellent prospects for home equity loan advertising(Grupe and Owrang 26).

4. Presentation of Information Obtained Through the Data Mining

The usual form of data mining results in tables of values. The human user should be able to interpret the gathered and sorted data and consider it usefulness for the business. There are a great number of visualization tools that assist the user in his or her decision making process. They include geographical maps, histograms, color coding, pie charts, tree-maps, three dimensional trends, scatter graphs, and heat maps. Visualization tools provide even more insights than rows and columns of a table. Visualization tools are interactive, meaning that the user can change and sort the information shown. These tools are far more advanced than the traditional graphic applications. For example, a five-dimensional chart can be created by representing clusters on a three-dimensional scatter chart as a sphere. The size and color of a sphere represent the fourth and fifth dimensions. The size and color of a sphere represent the fourth and fifth dimensions. The time dimension can be incorporated by playing it like a video. The user can watch the movements in a multi-dimensional chart as it changes with the elapsed time (Newing 1996).

Ⅳ. Online Analytical Processing(OLAP)

OLAP software can be used to identify trends, model complex relations among data elements via a multidimensional analysis of data from the data warehouse, data mart, or multidimensional databases. Today's extremely competitive business world, analytical capabilities of decision support tools are necessary to create breakthrough strategies and see previously hidden opportunities. Middle and lower level management need easy-to-use decision support tools that can be learned during several short sessions and that do not require extensive and in depth computer skills.

This process involves multidimensional data analysis. We are used to two dimensional data presentation. For example, a graph of past revenues is two dimensional(revenues and years). However, with all amounts of data we have, two dimensions might not be enough. OLAP allows to interact data and work with it simultaneously. OLAP can present data in several dimensions which makes it better for analysis purposes(Castelluccio 1996).

OLAP allows managers to go through many dimensions and levels of data. OLAP model can be perceived as a cube with variables(actual data) as different dimensions(Baum 1996). It can be looked as a Rubik's Cube with data that can be twisted and turned in various ways by a manager to perform sensitivity analysis or future forecast scenario.

OLAPs can be separated into three groups according to their analytical capabilities. The first group allows multidimensional viewing, result reporting, and very little analysis. The second group offers advanced analysis features like rolling averages and top ten selection. The third group gives a user to perform analysis with the help of algorithms, pattern analysis, and allocation rules(Darling 1996).

One of the most important characteristics of a good OLAP server is flexibility and ability to drill the cube, skipping several levels at a time and going in any direction that is necessary. Another must have is ability to do multi-user support and time-series analysis (The 1995).

The primary purpose of OLAP is to quickly and efficiently manipulate and create various combinations of ad hoc business data. "OLAP tools are not production transaction-processing database. While they are conceptually simple, they lack the mathematical foundation of the relational model and have not proven themselves in day-to-day, mission critical operations"(Varhol 1995). But this is not their requirement, they are created to manipulate data. Though the more advanced OLAPs have analytical capabilities.

OLAP becomes more and more popular since it allows managers to understand a number or related factors (greater than three) over a period of time. OLAP can answer a question like that "How much did the company spend on health benefits, by in division X, in each year, compared with the plan?"(Baum 1996).

Dominant types of OLAP databases pushed by vendors are ROLAP(relational online analytical processing) and MOLAP(multidimensional online analytical processing). Each form of multidimensional databases has its own advantages and disadvantages. For example, MOLAP is better when intense computation is frequently necessary and quicker response time is needed, whereas ROLAP is a better choice when dealing with maximum number of data dimensions and larger volume of data(See Gray 1996 for a more detailed comparison between the two).

Interpretation and Evaluation involve interpreting the discovered patterns /relationships /classes that can be understandable by user. The KDD is an on-going process begins from selecting target data(as shown in figure 1), preprocessing transforming data to find knowledge, and interpreting the patterns. The KDD definition as an on-going process includes a possibility of returning to any previous steps in the figure.

V. Summary and Conclusion

As reviewed, a host of new tools are emerging in the DSS area. These emerging tools include the data warehouse/multidimensional databases (MDDB), data mining, on-line analytical processing(OLAP), intelligent agents, World Wide Web technologies, the Internet, and corporate intranets. They are becoming an integral part of a set of recent developments in data management, DSS, expert systems, and EIS/ESS. These tools are reshaping DSS developments in organizations. These new tools are a set of inseparable tools that adds new capabilities to DSS, EIS, and ESS.

5.1. A set of emerging new tools such as the data warehouse, KDD, OLAP, etc. allow today's managers to make sound and effective decisions. Most traditional two dimensional relational databases with limited query ability can answer questions concerning historical information such as "How did we do last year in division X comparing with division Y?" or "What was our financial performance in the last ten years?" But managers are not satisfied with this kind of information. They have to make decisions concerning the future, so they need forecasting and predictive tools.

First, data warehouses accumulate and sort all the available data in the useful manner. Then data mining and OLAP applications allow managers to extract new knowledge that was previously hidden. They provide forecasting of the future by using artificial intelligence, neural network, decision trees, what-if techniques, genetic algorithms and scenario analysis. These tools allow managers to make better decisions to improve the operations of their businesses.

5.2. The emergence of Data Warehouse Decision Support Systems. A new type of DSS has been introduced in the market. For example, RiskAdvisor is a new knowledge-based DSS to help users (underwriters, claim examiners, financial analysts, auditors, etc.) in making 19 insurance industry functions such as premium analysis, loss analysis, accounts profiles, case-load management, etc.(Anonymous, 1996). The knowledge obtained through the database mining can be used in different ways by managers and executives to make better decisions. It analyzes trillions of transactions at once to figure out sales trends of a certain seasonal product such as ice cream, lemonade by the season, the months, the geographical region, etc. It can also spot a subtle shift in fashion trends for a certain product. It can also be used for marketing mix decisions.

5.3. The knowledge obtained through the database mining can be used in different ways by knowledge engineers for creation of new knowledge bases or for improving and updating the existing knowledge base. The experts and generated by the data mining tools will be capable of making better quality decisions. Furthermore, The knowledge obtained though the database mining can be used in different ways by database system managers to improve existing databases by improving the logical database design.

5.4. The data warehouse allows an easier access to useful information and speeds decision making process. Consequently, we will expect performance and efficiency improvement, which in turn will lead higher productivity and lower costs. Due the integrated nature of data in the data warehouse, decision makers focus on making decisions using the data, not cleansing the data. A data warehouse contains only data that is necessary for decisions support systems and executive support systems. A significant advantage of the warehouse data over the data from operational data bases is aggregated/summarized data.

5.5. Since information is provided on timely basis and external data such as market research, economic forecast, and competitor information can be integrated into the organization's internal database, and competitor information can be integrated into the organization's internal databases, business can react rapidly to changing environment.

Ⅵ. Implications and Discussion

All these developments we have reviewed suggest that we are shifting the corporate decision making paradigm from information-driven decisions in the 1980s to knowledge driven decisions in the 1990s. Expert decision support systems(Intelligent DSS) are incorporating domain knowledge and modeling and analysis system to provide users the capability of intelligent assistance. Knowledge base modules are being used to formulate problems and decision models, analyze and interpret the results. Some systems are adding knowledge-based modules to replace human judgements. Managerial judgements have been used to ascertain(assess)future uncertainty, and to select assumptions on which decision model can be based on. Some decisions are both knowledge and data intensive. Consequently, a large amount of data usually require considerable effort to for their interpretation and use.

Knowledge-based systems are effective tools for interpreting these data to diagnose the cause of unsatisfactory symptoms in the manufacturing processes for example. The need to deal with a large amount of uncertain data necessitates the development of a knowledge-based DSS.(e.g., Guida and others 1992).

A new set of DSS tools and approaches are adding a new set of capabilities of DSS to effectively deal with the problems that cannot be dealt with previously. The majority of these tools are being used to make decision support systems intelligent by providing a computer systems with a new ability. The new ability includes machine learning which refers to the computational methods/tools of a computer system to learn from experience (past solutions), data, observations, and consequently alter its behavior, triggered by modification to the stored knowledge. Artificial neural networks and genetic algorithms are the most notable approaches to machine learning. Other tools include fuzzy sets, modeling by example, geographical information system(GIS), logic modeling, and visual interactive modeling(VIM).

Perhaps the recent development of AI imbedded data management tools we have reviewed may portend that more and more MIS/DSS systems are progressively being developed to replace managerial judgment and making decisions with fewer human inputs in the form of intuitions and managerial judgment. If the trend of emerging knowledge DSS continues, this will open the possibility of wider application of DSS into the realm of strategic decision support, where a wide array of human judgement and knowledge is necessary. After a great deal of conceptual discussions in the DSS literature on so called active DSS, we begin to see prototype systems that show the active role of knowledge-based components toward fruitful collaboration between thee human and the machine in constructing linear programming models, analyzing econometric models, and automatically adjusting scheduling models in response to unanticipated events on the factory floors (Blanning and King 1991). The essence of active decision support activities includes monitoring decision making processes, stimulating creative ideas through carrying out insightful conversations with decision makers.

References

Abbort, Daniel, "Insurance companies launch high-tech assault on fraud," CPCU Journal, Fall 1996, Vol 49., No. 3, pp. 7-8.

Angehrn, A.A. "Computers that criticize you: Stimulus-based decision support Systems, "Interfaces, 1993, Vol. 23, No.3, pp. 3-16.

Anonymous, "Agency for Health Care Policy and Research Announces Project in Clinical Decision Support," http://www.med.uiucedu./ahcpr.html, 26 Jan., 1997.

Anonymous, "Platinum unveils data warehouse decision support system," Best's Review, Property-Casualty Insurance Edition, December 1, 1996, Vol. 96, No.8, p. 103.

Baum, David, Data Warehouse, Oracle, March/April 10, 1996, pp. 34-43.

Brachman, Ronald J., Khabaza, Tom, Kloesgen, Willi, Piatesky-Shapiro, Gregory, & Simoudis, Evangelos. "Mining Business Databases," Communications of the ACM, 1996, November, Vol. 39, 11, 42-48.

Brooks, Peter. "Data Mining Today," DBMS, 1996, October 10, pp. 59-64.

Castelluccio, Michael. "Data Warehouse, Marts, Metadata, OLAP/ROLAP, and Data Mining - a Glossary," Management Acco-unting, October, 1996, Vol. 78, pp. 59-61.

Chevinsky, Hannan and Henderson, Deborah (1997, February), "Building a Cost-Effective Data Warehouse," CMA Magazine, Feburary, 1997, Vol. 71, No. 1, pp. 19-23.

Codd, E, F., Codd, S.B., and Sally, C.T. "Beyond Decision Support," Computerworld, July 26, 1993, Vol. 27, No. 30, pp. 87-89.

Culpepper, Kenneth. "Database Maintenance Planning :House-Breaking the Monster," Direct Marketing, October, 1996, Vol. 59, pp. 44-60.

Darling, Charles B. "How to Integrate Your Data Warehouse," May 15, 1996, Datamation, Vol. 42, pp.40-51.

Darling, Charles B. "Think Outside the OLAP Box," Datamation, April 15, 1996, Vol. 42, pp. 88-92

Edenstein, Herb, & Millenson, Janet, Lessons From the Trenches : Knowledge, Discovery, and Data Mining," DBMS, February, 1996, Vol. 10, p. 62.

Eom. S. B. 1996a, "Global Management Support Systems: The New Frontiers," in Global Information Technology and Systems Management, eds. Palvia, P.C., Palvia, S.C., and Roche, E.C., Ivy League Publishing, Nashua, New Hampshire, pp. 441-460.

Fayyard Usama, Haussler, David, & Stolorz, Paul. "Mining Scientific Data," Communications of The ACM, November, 1996, Vol. 39, No. 11, pp. 27-34

Fayyard Usama and Gregory Piatetsky-Shapiro, and Padhraic Smyth, "The KDD Process for Extracting Useful Knowlege from Volumes of Data," Communications of the ACM, November 1996, Vol. 39, No. 11, pp. 27-34

Fayyard Usamma and Ramasamy Uthu-rusamy, "Data Mining and Knowledge Discovery in Databases : Introduction, "Communications of the ACM, November 1996, Vol. 39, No. 11, pp. 24-26.

Foley, John, "Data Dilemma," Information Week, June 10, 1996, pp. 14-16.

Glymour Clark, David Madigan, Daryl Pregibon, and Padhraic Smyth, "Statistical Inference and Data Mining," Communications of the ACM, November 1996, Vol. 39, No. 11, pp. 35-41.

Gray Paul, "Data Warehouse, OLAP, Data Mining, and the New EIS," Tutorial Presented at Association for Information Systems(AIS) '96 Meeting, Phoenix, AZ., 1996.

Grupe, Fritz H., & Owrang, Mendi M., "Discovering New Knowledge and Competitive Advantage," Information Systems, Fall 1995, Vol. 12, pp. 26-32.

Halper, Mark, "Setting Up is Hard to Do : Data Warehouses Do Not Grow on Trees," Forbes, April 9, 1996, Vol. 157, pp. 50-52.

Hedberg, Sara Reese, "The Data Gold Rush," Byte, October 1995, Vol. 20, pp. 83-88.

Imielinski Tomasz and Heikki Mannila, "A Database Perpective on Knowledge Discovery," Communications of the ACM, November 1996, Vol. 39, No. 11, pp. 58-64.

Inmon W. H., "Buliding the Data Warehouse," New York : John Wiley & Sons, 1993.

Inmon W. H. & Hackathorn, Richard D., "Using the Data Warehouse," New York : John Wiley & Sons, 1994.

Inmon W. H. and C. Kelley Rdb/VMS, "Developing the Data Warehouse," Wellesley, MA., QED Publishing Group, 1993.

Immon W. H., "The Data Warehouse and Data Mining," Communication of the ACM, November 1996, Vol. 39, No. 11, pp. 49-50

Kight, Betty. "The Smart Way to Build a Data Warehouse," Datamation, October 1996, Vol. 42, pp. 91-94.

Kimball R., "The Data Warehouse Toolkit : Practical Techniques for Building Dimensinal Data Warehouses," New York : John Wiley & Sons, 1996.

Masson, Barry, "Data Mining : Exploring the Unknown," Discount Merchandiser, October 1995, Vol. 35, T58 - T60.

McElreath, Jack, "An Arcthitectual Perspective of Data Warehouses," Information Strategy, Summer 1996, Vol. 12, No. 4, pp. 30- 40.

Newing, Rod, "Data Mining," Management Accounting, October 1996, Vol. 74, pp. 34 - 36.

Newing, Rod, "Data Warehousing," Management Accounting, March 1996, Vol. 74, pp. 21 - 23.

Ott, J. 1992, "Federal Express Develops C31-based Information System," Aviation Week & Space Technology, November 23, 1992, Vol. 137, No. 21, pp. 57 - 60.

Poe, Vidette, "Building a Data Warehouse for Decision Support," Printice Hall PTR, NJ., 1996.

Riecken Douglas, "Introduction to Special Issue on Intelligent Agents," Communications of the ACM, July 1994, Vol. 37, No. 7, pp. 18 - 21.

The, Lee, "OLAP Answers Tough Business Questions," Datamation, May 1, 1995, Vol. 41, pp. 65 - 67.

Varhol, Peter D., "Three Routs to OLAP," Datamation, August 15, Vol. 41, pp. 57 - 63.