

A HOLISTIC FRAMEWORK FOR KNOWLEDGE MANAGEMENT

Dr. Shamsul Chowdhury, Roosevelt University, schowdhu@roosevelt.edu

ABSTRACT

Knowledge management refers to the set of processes developed in an organization to create, gather, store, transfer, acquire and apply knowledge. Today's business world is rather complex. We need flexible tools to deal with the evolving complexities. Different, as well as dynamic, business solutions need to be derived to cope with the ever-changing business world. Business problems and solutions not only affect people within the organizations, but also other people like customers and suppliers. Knowledge management increases the ability of the organization to learn from its environment and to incorporate knowledge into the business processes by adapting to new tools and technologies. Data mining tools and technologies have evolved during the last decade and now are extensively used by different organizations. Still the reusability of the knowledge that is being extracted is not very common. One way to improve the reusability is to use the knowledge base as front-ends to Case-based Reasoning (CBR) applications. The paper focuses on the issues of knowledge management and presents a holistic framework for knowledge management by integrating Data Mining (DM) tools and technologies with CBR tools and technologies. The purpose of the integrated framework is to attain, retain, revise and reuse Business Intelligence (BI) within an organization. BI is the process of gathering and using information in the field of business. In other words it is the process of converting data into information and then into knowledge to solve business problems. Business intelligence is carried out to gain sustainable competitive advantage and is a valuable core competence in today's complex world.

Key words: Data mining, Knowledge base, Business intelligence, CBR system

INTRODUCTION

Knowledge management refers to the set of processes developed in an organization to create, gather, store, transfer, acquire and apply knowledge [10,11, 17, 18]. Today's business world is rather complex. We need flexible tools to deal with the evolving complexities. Different, as well as dynamic, business solutions need to be derived to cope with the ever-changing business world. Business problems and solutions not only affect people within the organizations, but also other people like customers and suppliers. Knowledge management increases the ability of the organization to learn from its environment and to incorporate knowledge into the business processes by adapting to new tools and technologies [6,7, 19].

Over the years database applications, data warehouses, and data mining tools have evolved into a unique and popular business solutions to attain business intelligence. Decision makers already consider these systems to be key components of their IT strategy and architecture. It is no secret, especially considering the global nature of today's economy, that business intelligence has won a great deal of attention from top practitioners and researchers.

BI is carried out to gain sustainable competitive advantage and is a valuable core competence in today's complex world. BI needs access to combined and centralized business information. Data warehousing is being used by many organizations as a way to pull all related data together in order to successfully answer relevant business questions. Idea behind a data warehouse is to centralize company wide information to create and deliver the necessary analytical environment (for data mining or knowledge discovery) to meet the changing needs of business. Data warehouse data can be compared with the metaphor of the crystal river with its pure life-giving water, which always reflects the stone over which it flows. Similarly data warehouse data would reflect the organizational activities [14]. Business must treat user data with the same care that physicians apply to patient information [2, 3, 8, 9, 21].

Data and knowledge management in the form of sharing information within the enterprise and other business partners is a key success factor. BI is all about knowing where the market is going, where do we stand as a competitive organization, where there are threats and potential opportunities. Business intelligence system can best be defined as a broad category of applications and technologies for gathering, storing, analyzing, and providing access to data to help enterprise users make better business decisions. BI applications include the activities of decision support systems, query and reporting, online analytical processing (OLAP), statistical analysis, forecasting and data mining [12, 22].

We propose that in addition to the existing category of applications BI system should also include CBR systems to incorporate the reusability of the knowledge gained over time and retained to solve future problems. We intend to present a framework for integrating CBR systems along with other BI category of applications [13]. The proposed framework would be helpful in implementing information systems that would meet the organizational needs and enhance decision making by attaining, retaining and reusing BI in the organization.

Data Mining

Current trends in today's complex business environment include collecting and analyzing information regarding customers, products, and processes in order to increase profit and reduce expenses. Businesses are attempting to gather and comb through this data by implementing various systems to make up a suite of Business Intelligence System. BI is the process of gathering information in the field of business. The purpose is to enhance data into information and then into knowledge to be able to address business issues and concern successfully. The process can be summarized in the following scheme:

Data ---> Information ---- > Understanding ---- > Knowledge ---- > Business Intelligence

From the scheme above it can be seen that the data could be analyzed and interpreted to gain an understanding and increase our knowledge, which could be applied to handle business situation at hand. For a successful interpretation/evaluation of a data material we need knowledge in three different contexts, namely: - data context, - statistical context, - domain (business) context [4].

The data context utilizes some general knowledge of the application area to transform or structure the data into a model, which could be statistical or analytical. The domain context, in

addition to the above, should have more deep and specific knowledge of the application area to be able to interpret the results appropriately and draw conclusions from the statistical analysis.

The data evaluation process described above is broadly known as Data mining or Knowledge extraction/discovery in databases (KDD) and could broadly be defined as the non-trivial extraction of implicit, previously unknown and potentially useful knowledge from data. The identified knowledge could be used to serve the following objectives: - make predictions or classifications about new data, - explain existing data, - summarize the contents of a large database to facilitate decision making. The most commonly used techniques in data mining today are [20]:

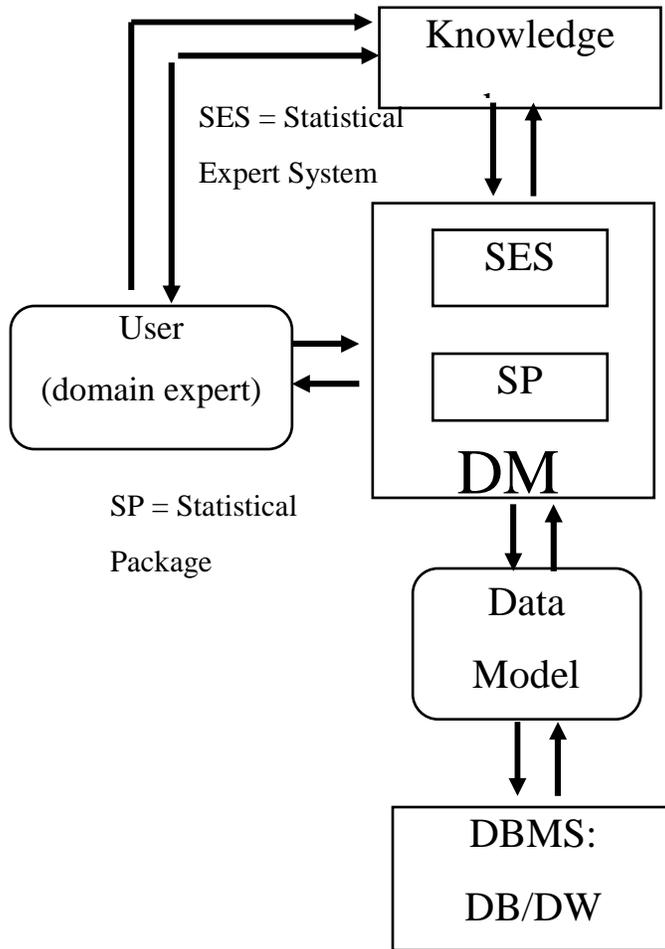


Figure 1. A General Framework for Data mining Methodology

- **Artificial neural networks (ANN):** Non-linear predictive models that learn through training and resemble biological neural networks in structure.
- **Decision trees:** Tree-shaped structures that represent sets of decisions. These decisions generate rules for the classification of a dataset.

- **Genetic algorithms:** Optimization techniques that use processes such as genetic combination, mutation, and natural selection in a design based on the concepts of evolution.
- **Nearest neighbor method:** A technique that classifies each record in a dataset based on a combination of the classes of the k record(s) most similar to it in a historical dataset. Sometimes called the k-nearest neighbor technique.
- **Rule induction:** The extraction of useful if-then rules from data based on statistical significance.

DM tools and technologies have evolved during the last decade and now are extensively used by different organizations. The strength of today's DM tools and technologies is due to – increasing computing power, improved data collection, and statistical and learning algorithm. A general methodology for data mining is given in figure 1. We implemented a system based on the above framework and it was proven successful in mining data from a large database described elsewhere [4-6]. Most data mining tools of today use a similar framework.

Case-based Reasoning

DM tools can help us extensively to learn and gain knowledge from the data sets. Still the reusability of the knowledge that is being extracted/gained is not very common. One way to improve the reusability is to use this knowledge base as front-ends to CBR applications. In other words DM tools and technologies need to be integrated with CBR tools and technologies to reuse the knowledge in addressing identical future problems. CBR is an artificial intelligence technology that uses previous knowledge from solved problems in order to give suggestions to the solution of new problems. In CBR, descriptions of previous expert experiences are represented as cases in a knowledge base, new problems are solved by adapting to solutions of similar cases in the past. The work of Schank and Abelson's in 1977 is widely held to be the origins of CBR [15,16, 24]. They proposed that our general knowledge about situations be recorded as scripts that allow us to set up expectations and perform inferences. The philosophical roots of CBR, is the work of Roger Schank's group at Yale University in the early 1980s, that produced both a cognitive model for CBR and the first CBR applications based upon this model. Janet Kolodner developed the first CBR system called CYRUS [15,16]. CYRUS contained knowledge, as cases, of the travels and meetings of ex-US Secretary-of-State Cyrus Vance. CYRUS was an implementation of Schank's dynamic memory model [24]. The CBR cycle or process is shown in Figure 2. The cycle can be described in short by using the four REs: - REtrieve the most similar case(s); - REuse the case(s) to attempt to solve the problem; - REvise the proposed solution if necessary; and RETain the new solution as part of a new case.

A new problem is solved by trying to describe what the problem is. This description is then used to retrieve old cases from the case database. The retrieved case is compared with the new case to reuse the previously gained knowledge. This process creates a suggested solution to the problem, which is compared with the real world example and tested. During this revision, useful experience from the test is added and forms a new tested solution. This tested solution gives a confirmed solution, which is then retained in the case database as a new case.

We have conducted some studies based on DM and CBR technologies and the results were presented elsewhere [5,6]. In these works DM and CBR were conducted separately and in phases on a large database to discover the potential causes of injuries (DM part- employing DM tools) and find ways of prevention of the injuries (CBR part – using CBR tool). The idea was to use the injury database (with causes of injuries) and transform it into a case base (causes of injuries + their preventions).

DM and CBR technologies can be combined together into a knowledge management holistic framework for attaining BI by performing data mining and retaining the BI by using CRR technology [6]. A framework for integration is presented in figure 3. System implemented based on the framework would help to acquire new knowledge through iterations of analysis and operations (data mining). Further the knowledge that is being discovered/captured by performing extensive analysis would be represented and reused by employing CBR technology to solve future problems. In other words the knowledge (discovered and stored by the DM framework – (Figure 3) in *knowledge base* would be used as a front-end by the CBR technology component (Figure 3) to create reusable CBR applications [5,6]. Currently we are exploring the possibilities of combining data mining technology together with CBR methodology based on the integrated framework to implement reusable systems for improved health care and better customer relationship management. We intend to present the results in our fourth coming works.

A Model for Case Based Reasoning

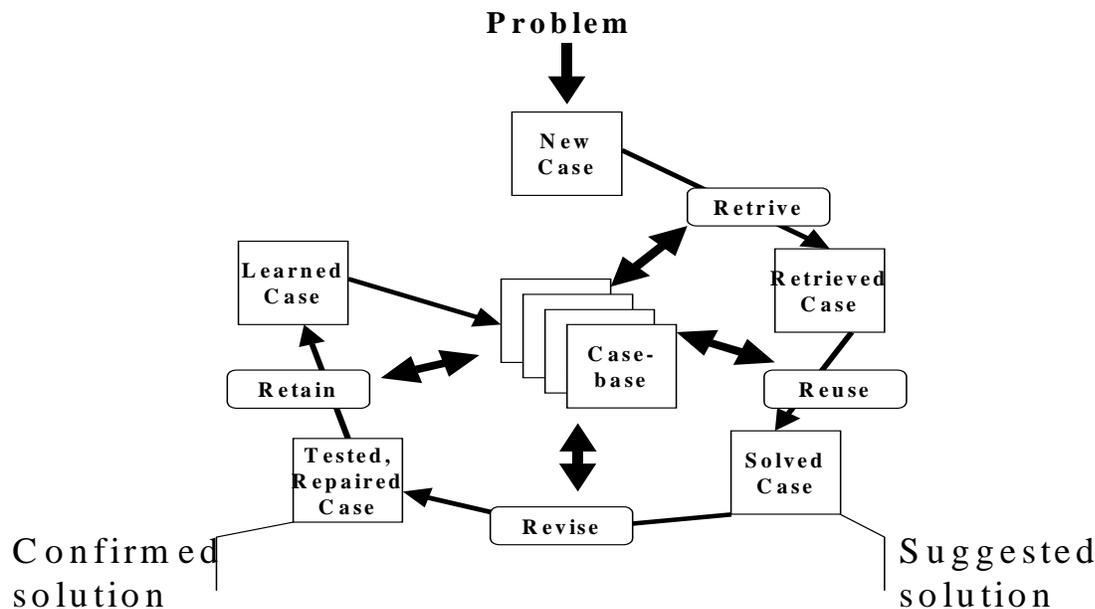


Figure 2. The CBR Cycle (Adopted from Aamodt & Plaza).

DISCUSSION

Data mining is used to support the acquisition of knowledge from databases. DM integrated with CBR (Figure 3) can be viewed as a general knowledge management and also a problem solving

tool. Knowledge Management is also the practice of treating knowledge as a corporate asset. Knowledge assets are equally important for competitive advantage and survival as physical and financial assets [3, 23]. Data mining and CBR can be seen as complimentary methodologies and integration of the two would help to implement powerful and reusable systems for better knowledge management as well as to address future problems. One of the key benefits of including CBR in the suite of business intelligence system would be to enhance organization's knowledge base by generating solutions to specific problems that are too massive and complex to be analyzed by human beings in a short period of time.

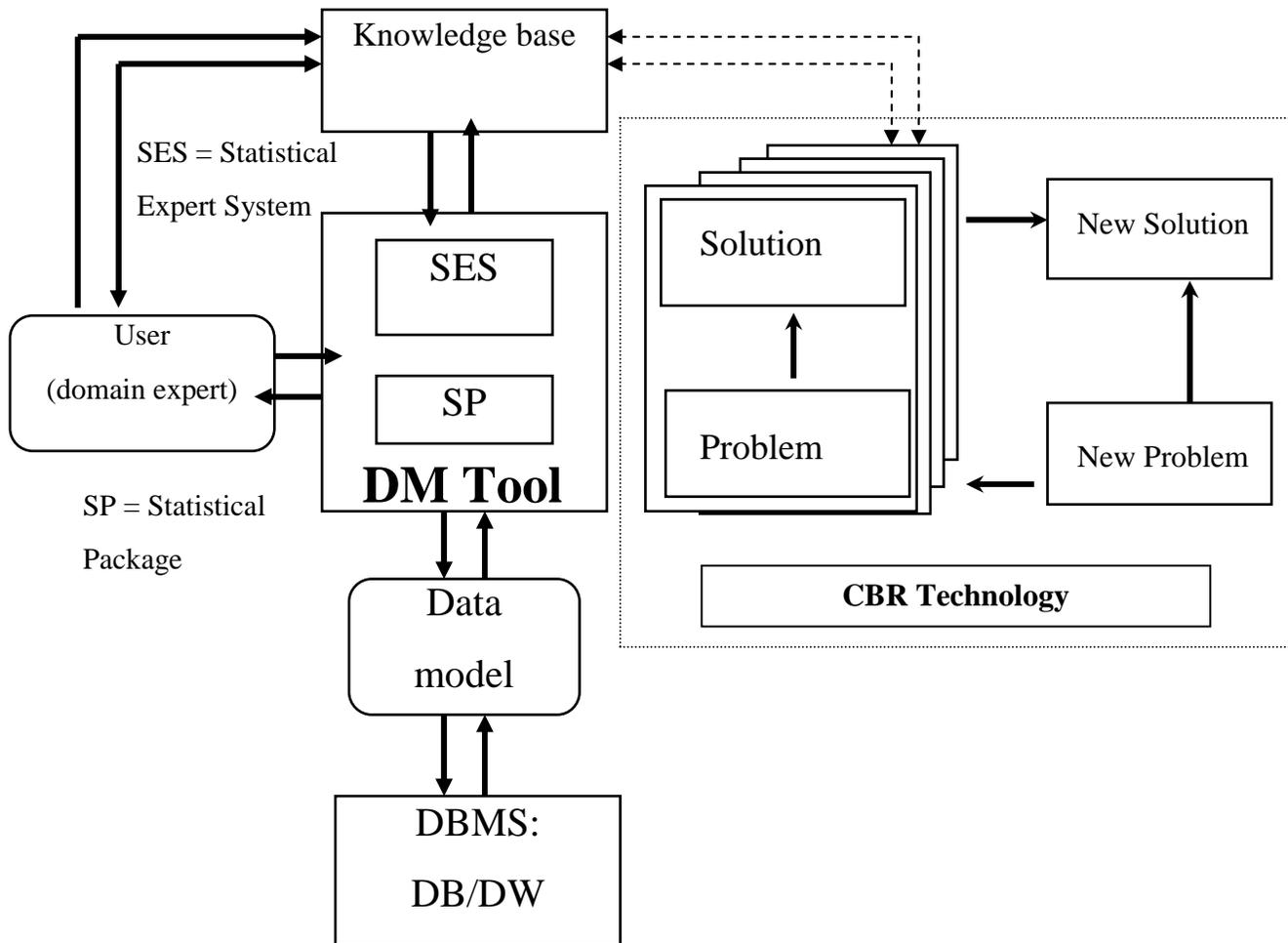


Figure 3. An Architecture for Integration of DM and CBR Technology

REFERENCES

1. Aamodt A. & Plaza E. (1995) Case-Based Reasoning – Foundation issues, methodological Variations, and System Approaches. *Artificial Intelligence Communications*, 7(1).

2. Albert, S. & Bradley, K. (1997). *Managing knowledge: Experts, agencies and organizations*, New York, NY: Cambridge University Press.
3. Carr, G. N. (2001). *The Digital Enterprise*, Boston, Ma: Harvard Business School Publishing Corporation.
4. Chowdhury, S.I. (1990). Computer-Based Support for Knowledge Extraction from Clinical Databases, Linköpings Studies in Science & Technology. *Dissertations* No. 240, Linköping University.
5. Chowdhury, S., Lindkvist, K. Åhlgren, M. & Timpka, T. (1998). Knowledge Discovery and Case-based Reasoning in Health Promotion: Development of a Help-Desk for Prevention of Occupational Injuries. In MEDINFO 98, B. Cesnik et al. (Eds). Amsterdam: IOS Press.
6. Chowdhury, S. (2001). From Databases to Case-bases: What, Why and How?" presented at the MBAA (Midwest Business Administration Association) –SAIS (Society for the Advancement of Information System) Conference.
7. Combs, E., Richard, D. & Moorhead, J. (1992). *The Competitive Intelligence: Handbook*. Metuchen, N.J: The Scarecrow Press, Inc.
8. Data Management Association. Guidelines to Implementing Data Resource Management. Bellevue, WA: DAMA International, 2002.
9. Drucker, F.P. (1999). *Management Challenges for the 21st Century*, New York, NY: Harper Collins Publishers, Inc
10. Harmon, Fred. (2001). *Business 2010*. Washington DC: The Kiplinger Washington Ed., Inc.
11. Honeycutt, Jerry. (2000). *Knowledge Management Strategies*. Redmond, Washington: Microsoft Press
12. Inmon, H., Imhoff, W. & Ryan, C. (2001). *Corporate Information*, New York, NY: John Wiley & Sons, Inc.
13. Jayaratna, N. (1994). *Understanding and Evaluating Methodologies - A Systemic Framework*, McGraw-Hill, London.
14. Kimball, R. (1996). *The Data Warehouse Toolkit: Practical Techniques for Building Dimensional Data Warehouses*, John Wiley.
15. Kolodner J. (1993). *Case-based Reasoning*, Morgan Kaufman Publishers: San Mateo.
16. Kolodner J. (1991). Improving Human Decision Making through Case-Based Division Aiding, *AI Magazine* 1991, 52-68.
17. Kroenke, D. (2002). *Database Processing – Fundamentals, Design and Implementation* (8th ed.), Pearson Education, Inc.
18. Laudon, K.C & Laudon, J.P. (2004). *Managing the Digital Firm*, (8th ed.), Upper Saddle River, New Jersey: Prentice Hall.
19. Liautaud, B. & Hammond M. (2001). *e-Business Intelligence: Turning Information into Knowledge into Profit*, McGraw-Hill
20. Mailvaganam, H. (nd). Models for Data Mining, *DM Review*, Retrieved 25 April 2005. <http://www.dwreview.com/Data_mining/DM_models.html>
21. Miller, P. (2000). Jerry and the Business Intelligence Braintrust. *Millennium Intelligence: Understanding and conducting Competitive Intelligence in the Digital Age*, Medford, New Jersey: Information Today, Inc.
22. Pendse, N. (nd). The OLAP Report. <http://www.olapreport.com/>
23. Russ, M. (2003). What is new about the new economy? *The Green Bay Area Chamber of Commerce Friday Report*, Nov. 21, 10(22).
24. Schank R.C. (1989). *Inside Case-Based Reasoning*, Hillsbury, NJ: Lawrence Erlbaum.