Organizational Data Mining

Hamid R. Nemati
The University of North Carolina at Greensboro, USA

Christopher D. Barko
The University of North Carolina at Greensboro, USA

INTRODUCTION

Data mining is now largely recognized as a business imperative and considered essential for enabling the execution of successful organizational strategies. The adoption rate of data mining by enterprises is growing quickly, due to noteworthy industry results in applications such as credit assessment, risk management, market segmentation, and the ever-increasing volumes of corporate data available for analysis. The quantity of data being captured is staggering—data experts estimate that in 2002, the world generated five exabytes of information. This amount of data is more than all the words ever spoken by human beings (Hardy, 2004). The rate of growth is just as astounding—the amount of data produced in 2002 was up 68% from just two years earlier. The size of the typical business database has grown a hundred-fold during the past five years as a result of Internet commerce, ever-expanding computer systems, and mandated recordkeeping by government regulations (Hardy, 2004). Following this trend, a recent survey of corporations across 23 countries revealed that the largest transactional database almost doubled in size to 18 TB (terabytes), while the largest decision-support database grew to almost 30 TB (Reddy, 2004).

However, in spite of this enormous growth in enterprise databases, research from IBM reveals that organizations use less than 1% of their data for analysis (Brown, 2002). In a similar study, a leading business intelligence firm surveyed executives at 450 companies and discovered that 90% of these organizations rely on gut instinct rather than hard facts for most of their decisions, because they lack the necessary information when they need it (Brown, 2002). In cases where sufficient business information is available, those organizations only are able to utilize less than 7% of it (The Economist, 2001). This is the fundamental irony of the Information Age we live in—organizations possess enormous amounts of business information yet have so little real business knowledge.

In the past, companies have struggled to make decisions because of lack of data. But in the current environment, more and more organizations are struggling to overcome information paralysis—there is so much data available that it is difficult to determine what is relevant and how to extract meaningful knowledge. Organizations today routinely collect and manage terabytes of data in their databases, thereby making information paralysis a key challenge in enterprise decision making. The generation and management of business data lose much of their potential organizational value unless important conclusions can be extracted from them quickly enough to influence decision making, while the business opportunity is still present. Managers must understand rapidly and thoroughly the factors driving their business in order to sustain a competitive advantage. Organizational speed and agility, supported by fact-based decision making, are critical to ensure an organization remains at least one step ahead of its competitors.

BACKGROUND

The manner in which organizations execute this intricate decision-making process is critical to their well-being and industry competitiveness. Those organizations making swift, fact-based decisions by optimally leveraging their data resources will outperform those organizations that do not. A robust technology that facilitates this process of optimal decision making is known as Organizational Data Mining (ODM). ODM is defined as leveraging data mining tools and technologies in order to enhance the decision-making process by transforming data into valuable and actionable knowledge to gain a competitive advantage (Nemati & Barko, 2001). ODM eliminates the guesswork that permeates so much of corporate decision making. By adopting ODM, an organization's managers and employees are able to act sooner rather than later, be proactive rather than reactive, and know rather than guess. ODM technology has helped many organizations to optimize internal resource allocations while better understanding and responding to the needs of their customers.

ODM spans a wide array of technologies, including, but not limited to, e-business intelligence, On-Line Analytical Processing (OLAP), optimization, Customer

Relationship Management (CRM), electronic CRM (e-CRM), Executive Information Systems (EIS), digital dashboards, and enterprise information portals. ODM enables organizations to answer questions about the past (what has happened?), the present (what is happening?), and the future (what might happen?). Armed with this capability, organizations can generate valuable knowledge from their data, which, in turn, enhances enterprise decisions. This decision-enhancing technology offers many advantages in operations (faster product development, optimal supply chain management), marketing (higher profitability and increased customer loyalty through more effective marketing campaigns), finance (optimal portfolio management, financial analytics), and strategy implementation (Business Performance Management [BPM] and the Balanced Scorecard).

Over the last three decades, the organizational role of information technology has evolved from efficiently processing large amounts of batch transactions to providing information in support of tactical and strategic decision-making activities. This evolution, from automating expensive manual systems to providing strategic organizational value, led to the birth of Decision Support Systems (DSS), such as data warehousing and data mining. The organizational need to combine data from multiple stand-alone systems (e.g., financial, manufacturing, and distribution) grew as corporations began to acknowledge the power of combining these data sources for reporting. This spurred the growth of data warehousing, where multiple data sources were stored in a format that supported advanced data analysis.

The slowness in adoption of ODM techniques in the 1990s was partly due to an organizational and cultural resistance. Business management always has been reluctant to trust something it does not fully understand. Until recently, most businesses were managed by instinct, intuition, and gut feeling. The transition over the past 20 years to a method of managing by the numbers is both the result of technology advances as well as a generational shift in the business world, as younger managers arrive with information technology training and experience.

ODM Research

Given the scarcity of past research in ODM along with its growing acceptance and importance in organizations, we conducted empirical research during the past several years that explored the utilization of ODM in organizations along with project implementation factors critical for success. We surveyed ODM professionals from multiple industries in both domestic and international organizations. Our initial research examined the ODM industry status and best practices, identified both technical and business issues related to ODM projects, and elaborated on how organizations are benefiting through enhanced enterprise decision making (Nemati & Barko, 2001). The results of our research suggest that ODM can improve the quality and accuracy of decisions for any organization that is willing to make the investment.

After exploring the status and utilization of ODM in organizations, we decided to focus subsequent research on how organizations implement ODM projects and on the factors critical to its success. To that end, we developed a new ODM Implementation Framework based on data, technology, organizations, and the Iron Triangle (Nemati & Barko, 2003). Our research demonstrated that selected organizational data mining project factors, when modeled under this new framework, have a significant influence on the successful implementation of ODM projects.

Given the promise of strengthening customer relationships and enhancing profits, CRM technology and associated research are gaining greater acceptance within organizations. However, findings from recent studies suggest that organizations generally fail to support their CRM efforts with complete data (Brohman et al., 2003). As further investigation, our latest research has focused on a specific ODM technology known as Electronic Customer Relationship Management (e-CRM) and its data integration role within organizations. Consequently, we developed a new e-CRM Value Framework to better examine the significance of integrating data from all customer touch-points with the goal of improving customer relationships and creating additional value for the firm. Our research findings suggest that, despite the cost and complexity, data integration for e-CRM projects contributes to a better understanding of the customer and leads to higher return on investment (ROI), a greater number of benefits, improved user satisfaction, and a higher probability of attaining a competitive advantage (Nemati, Barko & Moosa, 2003).

MAIN THRUST

Data mining is the process of discovering and interpreting previously unknown patterns in databases. It is a powerful technology that converts data into information and potentially actionable knowledge. However, there are many obstacles to the broad inclusion of data mining in organizations. Obtaining new knowledge in an organizational vacuum does not facilitate optimal decision making in a business setting. Simply incorporating data mining into the enterprise mix without considering non-technical issues is usually a recipe for failure. Businesses must give careful thought when weaving data mining into their organization's fabric. The unique orga-