## **Preface**

Mountains of business data are piling up in organizations every day. These organizations collect data from multiple sources, both internal and external. These sources include legacy systems, customer relationship management and enterprise resource planning applications, online and e-commerce systems, government organizations and business suppliers and partners. A recent study from the University of California at Berkeley found the amount of data organizations collect and store in enterprise databases doubles every year, and slightly more than half of this data will consist of "reference information," which is the kind of information strategic business applications and decision support systems demand (Kestelyn, 2002). Terabyte-sized (1,000 megabytes) databases are commonplace in organizations today, and this enormous growth will make petabyte-sized databases (1,000 terabytes) a reality within the next few years (Whiting, 2002). By 2004 the Gartner Group estimates worldwide data volumes will be 30 times those of 1999, which translates into more data having been produced in the last 30 years than during the previous 5,000 (Wurman, 1989).

This proclamation about data volume growth is no longer surprising, but continues to astound. Although for businesses, more data isn't always better. Organizations must assess what data they need to collect and how best to leverage it. Collecting, storing and managing business data and associated databases can be costly, and expending scarce resources to acquire extraneous data fuels inefficiency and hinders optimal performance. Managers must thoroughly understand the factors driving their business in order to optimize their data management efforts.

In spite of this enormous growth in enterprise databases, research from IBM revealed that organizations use less than 1 percent of their data for analysis (Brown, 2002). This is the fundamental irony of the "Information Age" we live in: Organizations possess an enormous amount of business information, yet have so little real business knowledge. And, to magnify the problem further, a leading business intelligence firm recently surveyed executives at 450 companies and discovered that 90 percent 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).

The solution for these problems lies in a technology known as Organizational Data Mining (ODM). ODM is defined as leveraging data mining tools and technologies 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. Advances in ODM technology have helped many organizations optimize internal resource allocations, while better understanding and responding to the needs of their customers.

ODM can analyze enormous volumes of data while determining the most pertinent data to focus on. The end result is optimal resource allocation and improved business performance.

This book is the result of numerous observations that data mining is gaining greater acceptance within organizations and delivering colossal benefits, yet there are few scholarly books devoted to the exploration of both the organizational and technical factors involved in leveraging enterprise data resources for optimal performance. Our objective is not to address all aspects of ODM — this would obviously require many more volumes. In addition, this book is not a data mining how-to manuscript — there are many of those available that adequately cover the technical aspects of data mining. Instead, we intend to explore and bring to light a number of interesting and practical domains within ODM, which integrates both data mining and organizational disciplines, such as organizational learning and behavior, business culture and ethics, business strategy and knowledge management. Practitioners, educators and members of the research and development community will each find relevant and enticing material in the following chapters.

## **ORGANIZATION OF THE BOOK**

Chapter I begins with an introduction to ODM by explaining what it is, how ODM is distinctly different from simple data mining, and briefly elaborate on its significance in today's competitive organizations. In addition, we take a look at the current status of ODM research, examine the evolution of ODM to the present day and contemplate its challenging yet opportunistic future. The remaining sections of the book explore ODM from different organizational perspectives. These sections are the Strategic Implications of ODM, Business Process Innovations through ODM, ODM Analytics and Algorithms, Industrial ODM Applications, and ODM Challenges and Opportunities. These sections contain many insightful and cutting-edge chapters from some of the best and brightest practitioners and researchers in ODM. We will now elaborate on each section's contents in more detail.

The first section, Strategic Implications of ODM, contains Chapters I through V on ODM's supporting role in formulating and implementing corporate strategy. Chapter II looks at multinational corporate sustainability from a content-analysis approach. The authors argue that companies are no longer motivated by finances and that societal concerns are being recognized and addressed by corporate board members of Fortune's Global 500. This chapter employs the use of content analysis of corporate messages to quantify the frequency/presence of certain words and phrases related to corporate sustainability within the annual reports of 24 multinational firms. Chapter III utilizes Michael Porter's Five Forces Model to understand the potential strategic value of ODM within the Australian Banking industry. This chapter explores how ODM can affect industry structure and attractiveness by assisting businesses, such as banks, to defend themselves against forces, such as those asserted by buyers, substitute products, new entrants and suppliers. Chapter IV investigates the role of ODM in supporting strategic decision making by providing validation for Micro-Theories (MTs), which are beliefs regarding the organization's task environment, such as sales increasing in a certain segment or customers preferring a certain product. This chapter suggests a four-step process for identifying and verifying MTs and illustrates this with a hypothetical example of a bank. Chapter V focuses on the privacy implications of ODM. The authors look at ethical data management issues from both an organizational and governmental viewpoint.

Section Two, Business Process Innovations through ODM, contains chapters exploring how ODM enables business process innovation. Chapter VI looks at the potential of knowledge exchange in organizations and provides readers with an understanding of the human dynamics of expert knowledge exchange in the realm of virtual teams. The authors present research, theory and the methodologies now in professional use to assess information exchange potential for knowledge management (KM) related activities at the team level as well as from the perspective of organizational culture. Chapter VII introduces practical issues of information navigation and organizational knowledge management involved in delivering customer service via the Internet. This chapter presents an adaptive, organic approach to knowledge management and discusses a state-of-the-art application named RightNow eService Center that embodies this concept. Chapter VIII discusses formulating a successful purchasing negotiation strategy and proposes that the buyer's level of expertise and/or simulated negotiation experiences through the experiential learning process help him/her better prepare for the negotiation and, thereby, increase his/her bargaining strength. This chapter uses both statistical data analysis and data mining techniques to demonstrate their usefulness in the optimal performance of business-to-business negotiations. Chapter IX focuses on extracting value from virtual organizational discussions through textual data mining. This chapter proposes a method of recovering value from the text of virtual group discussions based on methods derived from the communication field.

The third section, ODM Analytics and Algorithms, includes chapters of a more technical flavor and investigates new methods in ODM analytics and algorithms. Chapter X presents an integrated model in which data mining and online analytical processing (OLAP) complement each other to support intelligent decision making for data-rich environments. Chapter XI presents knowledge mining in decision support system model analysis. The three stages of mathematical modeling include model formulation, solution and analysis, with the later being seldomly addressed and, thus, the focus of this chapter. Chapter XII discusses recent advances in the use of agent technology in Decision Support Systems (DSS) and introduces a model for an agent-based DSS. This chapter informs the readers about the state-of-the-art in agent-based DSS and illustrates an example of an agent-based DSS for investment decisions. Chapter XIII proposes a methodology to scan, analyze and classify the content of primarily text-based Web documents to aid an organization in gathering information. The purpose of this chapter is to develop and demonstrate a methodology used to aid an organization in its environmental scanning efforts in light of the vast quantities of information available via the Internet.

The fourth section, Industrial ODM Applications, contains case studies of commercial applications of ODM and demonstrates how leading organizations are leveraging this technology for optimal performance. Chapter XIV discusses the basic concepts of data warehousing and, then, illustrates them through a case study of the global enterprise data warehouse (GEDW) implemented at 3M. The lessons learned at 3M can help other companies with their data warehousing initiatives. Chapter XV addresses ODM in franchise organizations. This chapter describes a comprehensive franchise framework, identifies the most important aspects of a franchising business and describes the role OLAP and data mining play and the necessary data items to ensure success. Chapter XVI examines the use of fuzzy clustering and expert reasoning for the identification of firms whose financial statements are affected by fraudulent financial reporting. The authors use a combination of fuzzy logic, expert reasoning and a statistical tool as an innovative method to evaluate the risk of fraudulent financial reporting. Chapter XVII describes ways in which library and information managers can use data mining in their libraries, i.e., bibliomining, to understand patterns of behavior among library users and staff members, and patterns of information resource use throughout the institution. This chapter presents a global view of the data generated in libraries and the variety of decisions that those data can inform. Chapter XVIII discusses how advances in data mining translate into business context. This chapter focuses on the retail industry and highlights the art of business implementation rather than the science of knowledge discovery in databases (KDD).

The last section, ODM Challenges and Opportunities, investigates the current challenges and opportunities surrounding ODM and how organizations are leveraging it to build or sustain a competitive advantage. Chapter XIX discusses impediments to exploratory ODM success. These impediments are based on an expert's anecdotal observations from multiple projects, either reviewed or undertaken by the author. The intent of the chapter is to provide an organization with a structure to anticipate these problems and prevent their occurrence in ODM efforts. Chapter XX argues that knowledge resides in human minds and is created by the continuous action and interaction happening in specific social contexts. This chapter highlights the advantages of adopting a constructionist knowledge approach and presents some constructionist guidelines to assist in the definition of ODM practices that leverage knowledge creation in organizations. Chapter XXI discusses the three phases of Web mining (data gathering, analysis and reporting) and describes each of these phases in detail along with a discussion of Electronic Customer Relationship Management (eCRM). In addition, integration issues and challenges surrounding Web mining and eCRM are explored. And lastly, Chapter XXII introduces a framework for ODM suited for both data-driven and hypotheses-driven problems. This ODM framework includes a comprehensive processing scheme that aims at increasing the benefits of ODM and other data analysis approaches by allowing a wider range of business problems to be tackled and by providing the users with structured guidance for planning and running analyses.

In closing, we believe that all enterprises are slowly moving from the Information Age to the Knowledge Age, where organizations will use ODM and supporting technologies to augment intuition with the purpose of allocating scarce enterprise resources for optimal performance. Industry professionals have suggested that many corporations could maintain current revenues at half the current costs, if they optimized their use of corporate data. Whether this finding is true or not, it sheds light on an important issue. Leading corporations in the next decade will adopt and weave these ODM technologies into the fabric of their organizations at all levels, from upper management all the way down to the lowest organizational level. Once the first organization within an industry realizes a competitive advantage through ODM and assimilates it into its decision-making process, it is only a matter of time before one of three events transpires: its industry competitors either adopt ODM, change industries or vanish. By adopting ODM, an organization's managers and employees are able to act sooner rather than later, anticipate rather than react, know rather than guess and, ultimately, succeed rather than fail.

## REFERENCES

Brown, E. (2002, April 1). Analyze this. Forbes, 169(8), 96-98.

- Kestelyn, J. (2002, August 12). No longer an afterthought. *Intelligent Enterprise*, 5(13), 6.
- Nemati, H. R., & Barko, C. D. (2001) Issues in organizational data mining: A survey of current practices. *Journal of Data Warehousing*, 6(1), 25-36.

Whiting, R. (2002, February 11). Tower of power. Information Week, (875), 40-43.

Wurman, R. (1989). *Information anxiety is produced by the ever-widening gap* between what we understand and what we think we should understand. New York: Doubleday.