Preface

Data mining can be termed, as the *extraction process of hidden predictive information from huge databases*. In a sense, *data mining* is a powerful emerging technology with enormous potential to assist communities focus on the most important information in their livelihoods. Utilizing data mining tools trends, patterns and behaviors of data can be monitored, and at the same time assist to predict future trends and behaviors of data. This allows societies/entrepreneurs/agencies to make proactive, intelligent, and knowledge-driven decisions. Furthermore, the automated and prospective analyses offered by data mining budge beyond the analyses of past events provided by demonstrative tools typical of decision support systems. Foremost, data mining tools can respond to diversified questions that traditionally were too time consuming to resolve. They scour databases for hidden patterns, finding predictive information that experts may miss because they lies outside their normal expectations (DIG White Paper, 1995).

Nowadays, data mining can also be termed as knowledge discovery processes, and increasingly becoming one of the most acknowledged topics in information technology. Much of it deals with the process of automatically retrieving essential information and discovering hidden relationships that is crucial in making knowledgeable decisions (Chen & Liu, 2005). During the last decade, data mining has been used in diversified application domains and has proven the ability to support various types of critical applications under knowledge discovery.

As an emerging field, data mining attracted potential researchers in this community, and at the same time, many have endorsed its diversified utilizations. Among several trends of its utilizations, archival of high-dimensional data; incorporation of semantic search; rapid data-driven analysis; knowledge-driven end products are phenomenal. In these contexts, social and political implications of data mining are important, especially in the realm of knowledge management in e-government. Ranging from providing value added social services, knowledge acquisition processes to electronic government initiatives, and content management systems implication of data mining converges. However, by far political implications of data mining in the knowledge management processes remain confined in the central bodies and corporate agencies. For proper knowledge development at the grass roots and for homogeneous distribution of knowledge acquisition system, bi-directional thoroughfare is desired in terms of policy initiations.

As Chaterjee (2005) outlines, data mining techniques are the outcome of an extensive process of study, research and product development. In essence, this evolution began when entrepreneurs started archiving business data in computers, efforts continued with improvements in easier data access, and more recently, researches generated technologies that allow users to navigate through their data in real time. Over the time, data mining takes this evolutionary development beyond retrospective data access, thus routing the expansion towards proactive information delivery for the management of human livelihood through improved intelligentsia.
Eventually, data mining is becoming a significant tool in science, engineering, industrial processes, healthcare, medicine, and other social services for making intelligent decision. However, the datasets in these fields are predominantly large, complex, and often noisy. Therefore, extracting knowledge from data requires the use of sophisticated, high-performance and principled analysis techniques and algorithms, based on sound statistical foundations. These techniques in turn necessitate powerful visualization technologies; implementations that must be carefully tuned for performance; software systems that are usable by scientists, engineers, and physicians as well as researchers; and infrastructures that support them (Mueller, 2004; SIAM, 2006).

Well developed information and communication network infrastructure and knowledge building applications, adapted to regional, national and local conditions, easily accessible and affordable, and making greater use of locally available technologies and other innovative methods where possible, can accelerate the social and economic progress of countries, and the well being of all individuals and communities. However, knowledge acquisition from crude data mostly depends on social, cultural, economical and political aspects of a country, rather than pure technology. It requires incorporation of intricately built knowledge management tools and techniques, archival of discrete data or information to feed the knowledge acquisition system and management of very superficial and complicated data or information within the system.

The issues and aspects of archiving scientific data include the discipline specific needs and practices of scientific communities as well as interdisciplinary values and methods. However, when this archival process is targeted to improve the livelihood of common people of the community, increase their knowledge level and eventually enhance the governance system, the tasks of data archiving ranges from accumulating, digitizing, preserving, providing security to providing easy accessibility. Thus, data archiving amalgamates with social diversifications, economic variations, cultural conjunctures, political peripheries and foremost, deals with human psychology. Data archiving becomes part and parcel of day-to-day activities of human life and management of these data reflects variation of human endeavors.

The leveraging of data mining tools to facilitate knowledge management and e-governance incorporates contexts of health care, literacy, governance issues, civic responsiveness, environment issues, climate changes and most importantly, equitable access for economic opportunities by creating social capital that are essential for the holistic deployment of economic development activities of a society. These issues are intricately intermingled with social and political environment of a nation. This book has tried to include issues, aspects, features, experiments, projects and cases related to socio-political enhancements among connected communities utilizing data mining techniques in the context of knowledge discovery and knowledge enhancement processes.

WHERE THIS BOOK STANDS

In this interconnected world, where information is an essential element of community empowerment, capacity development and socio-economic development, data mining is widening emerging and relatively a new area of research and development, which can provide important outcomes to the common citizens. It can yield substantial knowledge from even raw data that are primarily gathered for wider range of applications. In this context, various institutions have already started to derive considerable benefits from its diversified applications and at the same time, many other industries and disciplines are applying the technique in increasing effect for their own development.

However, the challenge of equitable distribution of resources for homogenous development taking into account the social, cultural, political and economic background protecting everyone’s interest confronts
policy initiators, academics and researchers, including government and non-government agencies, and development partners for taking attempt to create a development dynamics. Data mining would assist this evolutionary process beyond retrospective data access and routing to potential and proactive information delivery. This book includes various aspects of existing data mining and decision support techniques and their requirements to improvise the development of such information system at each tier of the governance structure. It also incorporated global initiatives and highlight their capabilities depicting real life cases to act as role models. The cases are on improvement of governance system, enhancement of peace building effort, electoral up-gradation, improvement of learning mechanism, scientific perfection, interactive imagery, or simple content repository for knowledge building.

Primary objective of this publication is to promote knowledge management in e-government, identify knowledge management technologies, and highlight the challenges in the implementation of e-government systems and especially knowledge management solutions at the grass roots. Moreover, as the contemporary world has began to realize that the political and economic significance of the more than half of the world’s population are living in largely untapped rural setting, realization of pragmatic knowledge acquisition processes to improve e-governance has become prime importance for every nation. For this reason, Governments and nongovernmental organizations, including development agencies are increasingly concerned with addressing economic development goals and stability, stubborn deficit in rural health and learning, urban migration, environmental degradation, and other related trends. In these circumstances, a publication on knowledge management processes focusing e-governance featuring emerging technology like data mining deserves primary importance among the academia, researchers, development practitioners, policy initiators and individuals.

Today many research communities, including informatics, systematics, academics, phylogenetics, ontologists, physicians, scientists and ecologists, are on the verge of transformation as a consequence of network capability that enable a new class of research dimension based on instant interaction with networked information servers and computational services, in a newly evolved paradigm known as ‘grid computing.’ High-throughput information backbones linking data enabled desktops of scientists around the globe assist to launch a new class of research collaboration with ‘wired together’ in a virtual and global environment. These processes are generating remarkable amount of data/content to be capitulated for effective deployment of knowledge acquisition system, and, therefore demand competent data/info mining techniques at the both end of usage (generator and end user). The book included conceptual framework, various analysis and synthesizes a variety of data mining techniques for improvement of social and political systems, varying from healthcare services, bio-energy database, crime analysis, and aviation safety to learning, portal management system, e-government modeling, and ontology.

The major asset of this book is the accumulation of several theoretical researches, case analysis, and a few practical implementation processes accompanying profound discussions and techniques for accomplishing tasks that one could easily adopt even in a non-technical environment. In fact, this book will act, not only as a research guide but also as an implementation guide in the longer run.

**ORGANIZATION OF THE BOOK**

The book is organized into five sections with seventeen chapters. A brief description of each of the chapters follows:

In recent years data warehousing and data mining are being used as unique tools in decision support, business intelligence and other kind of social and economic applications. **Chapter 1** of this book presents various data warehousing methodologies along with the main components of data mining tools
and technologies, and discusses how they all could be integrated together for knowledge management in a broader sense. The chapter also focuses on how data mining tools and technologies could be used in extracting knowledge from large databases or data warehouses. The chapter further focuses on the reusability issues of knowledge management and presents an integrated framework for knowledge management by combining data mining (DM) tools and technologies with case-based reasoning (CBR) methodologies.

While current literature on knowledge management has been addressing issues, challenges and opportunities for the private sector initiatives, little has been discussed for the public sector, and even less in the aspect of e-government. Acknowledging the fact that the implementation of a framework for the application of knowledge management in e-government is a challenging task that requires coordination and contribution from many agencies, departments and policy makers, Chapter II tried to put forward a framework for knowledge management in e-government for better e-government management. Furthermore, recognizing the importance of e-government and knowledge management (KM) to devolve into the public administration sector, the chapter put forward the continuation of a previous research related to the application of knowledge management in e-government. This chapter discusses important issues of KM in e-government and presented the framework for application of KM in e-government as a basis for future research. Finally, by positioning the proposed framework of KM in e-government, the chapter discusses about its application in a Greek municipality.

It is evident that appropriate data mining models and applications in e-government settings have the potential to bring major benefits to wider range of stakeholders. Furthermore, as these models evolve, structural transitions occur within e-government that includes an evolution of managerial practices through knowledge management (KM). Chapter III examines e-government challenges regarding the linkages between data mining and KM over time, discusses the organizational development of e-government applications, and details out both general and specific issues, as such social, ethical, legislative, and legal issues that impact effective implementations of e-government. A final focus of the chapter is the potential strategic benefits of a risk-based approach that can be used to enhance the core synergy of KM and data mining applications in e-government operations.

Knowledge management is not just a simple technology driven process, rather it is policy driven issue that is intermingled with technology, decision, management and intellectuality. Along this way, empowering common citizens utilizing knowledge development utilities is a challenge to the researchers and development practitioners. Furthermore, dissemination of intellectual content for public view, their understanding, capacity development, and specifically for being utilized as a tool to increase their social, educational, political and economic ability is by far the most difficult part of the system. Chapter IV of the book is focusing on knowledge management issues for developing knowledge management portals to empower citizens and societies. In this context, the chapter introduced a few critical aspects of knowledge management perceptions, justified establishment of knowledge management portals acting as a tool of empowerment, provided insight on data mining as a technology of implementation, suggested a solution by introducing Semantic Web Technologies as an essential technology for establishing knowledge management portals, puts forward contemporary challenges during the establishment of knowledge management portal, illustrated a few cases that are acting as knowledge management portals, and concluded before giving a few hints on future research issues for empowering common element of the society.

Utilizing data mining techniques, knowledge management in e-government is enhanced that evidently augments provision of improved social services by the government and other actors in the society. Along this perspective, the importance of various social services including education and health is constantly increasing. Simultaneously, as the societies are growing older, the awareness of a proper social care is
spreading among citizens. However, the central issue, here is the increasing cost of maintaining such social care systems, especially the expenses in healthcare are an important portion of the overall expenses of a country. Therefore, it is very important to determine if the given cares are the appropriate ones. **Chapter V** of the book is dealing with a methodology, *Health Discoverer*, and consequential software, aiming at disease management and measurement of appropriateness of cares. This chapter particularly focuses on data mining techniques that can be used to verify Clinical Practice Guidelines (CPGs) compliance and the discovery of new, better guidelines. The work is based on Quality Records, episode parsing using Ontologies and Hidden Markov Models.

As discussed in Chapter five, governance in healthcare refers to the complexity of checks and balances that determine how decisions are made within the top structures of hospitals. In this aspect, **Chapter VI** of this book introduces hospital governance as a policy domain in which data mining methods have a potential to provide insight and practical knowledge. The chapter starts by exploring the essentials of the concept, by analyzing the root notion of governance and comparing it with applications in other sectors of social services. The chapter also outlined the recent developments and examples in this sector from the UK, France and The Netherlands. Furthermore, a research agenda has been developed based on an evaluation of the current state of affairs. The chapter concludes with an introduction to the European Hospital Governance Project, which follows the outlines of the described research agenda.

Along the context of universal access to social services, apart from education and healthcare improved livelihood is important for better governance. **Chapter VII** focuses on data mining techniques that are being used to extract large quantity of data to get useful information. In this chapter the data mining approach is proposed for the characterization of family consumptions in Italy. The research has found that Italian expenditures form a complex system, and to mitigate this, the Italian National Bureau of Statistics (ISTAT) carries out a survey each year on the expenditure behaviour of Italian families. The survey enumerates household expenditures on durable and daily goods and on various services. For this purpose, a series of statistical techniques are used in sequence and different potentialities of selected methods for addressing these kinds of issues are pinpointed. This study recommends that, further investigation is needed to properly focalize on service usage for the characterization, for example, of the nature of investigated services (private or public) and, most of all, about their supply and effectiveness across the national territory. Still this study may be considered an example of operational and concrete approach of managing of large data-sets in the social-economical science, from the definition of goals to the evaluation of results.

Apart from education, healthcare, and livelihood services stable economic transactions is an important component of the governance system of a country. In this aspect, payment processes are the core of the economic transactions, whether the transaction is national or global. Payment system is an age-old system of transfer of properties. It has taken diverse forms depending on demand, usage, acceptability, tradition, culture, methods, technology and availability. Recently evolved payment systems varied from commodity transfers, physical financial transactions (traditional payment systems) to virtual payment transactions (electronic payment systems). Utilizing modern day technologies, electronic payment systems have also taken various forms in varying environments and societies. **Chapter VIII** provides a general overview on electronic payment systems, focusing developing countries and tried to related electronic payment systems as an enabler of financial empowerment. In this context, this chapter reiterated that by raising economic activities via electronic means, as a component of e-commerce, could augment the electronic governance system of a country. It has also put forward available issues, challenges, methods and tools needed to implement electronic payment systems, especially focusing developing countries.

**Chapter IX** discusses how a public sector organization made an effort to enhance their collaborative processes and organizational knowledge base in a specific functional area and how they implemented
KM-tools to improve knowledge and information sharing and transferal both within the organization and among partners. The implementation of KM-tools in the crime analysis function as part of intelligence-led policing assisted in preventing more crimes and creating safer communities. The use and recognition of the benefits of KM-tools also facilitated the development of continuing learning and education in all the levels and in all the fields and locations of the organization. This chapter aims to promote knowledge management and KM-tools usage in government and to identify the challenges and benefits gained by an implementation of analytical solutions. It also describes the managerial and organizational implications of the usage of analytical solutions and the effects on political decision-making and government programs.

In the arena of information security and knowledge management, text mining has becoming instrumental in extracting pertinent information to create valuable knowledge for more effective knowledge management. Chapter X addresses various technological applications of text mining in information systems security (ISS) issues. The techniques are being categorized according to the types of knowledge to be discovered and the text formats to be analyzed. The chapter also focuses on privacy issues of text mining, which are social aspects of text mining, and also discusses on future trends of text mining research.

Security, safety and trust are important components of a governance system, and as long as they are being managed by electronic means, they form quite a potential platform of electronic governance system. Along these issues, Chapter XI establishes aviation safety as a social service, forming part of e-governance and introduces a data analysis as an important application area for data mining. In this chapter, the reader is introduced to the basic concepts of data mining. After that, the field of aviation safety management is discussed, and in that connection, data mining is identified as a key technology to study through flight incidents reports. Afterwards the test runs for four data mining products, for probable use in the Finnish civil aviation authority, are described in detail. The chapter ends with conclusions that tell that even sophisticated data mining tools are just tools: they do not provide any automatic tools, but skilled users can use them for searching clues in the data.

It is apparent that, electronic government or digital government is not a simple or well-defined theoretical concept; rather it is a complex phenomenon, which involves technical, organizational, institutional and environmental aspects. Researchers from different disciplines are trying to model the E-government using combinations of various methods from different areas, which can help to deal with complexity and obtain explanations that are more comprehensive. Chapter XII uses Dynamics Causal Mining as the technique for modeling and analyze E-government, where Dynamics Causal Mining is a combination of System Dynamics and Data Mining. This chapter suggests an integration of System Dynamics and Association Mining for identifying causality and expanding the application area of both techniques. This gives an improved description of the target system represented by a database; it can also improve strategy selection and other forms of decision making. The aim is to identify causal factors hidden in the data and discover the underlying causality between the observed data.

In an effort to follow the new public administration roadmap and invest in the sharing of knowledge, most of the governmental organizations appear in a crossroad. A lot of knowledge has been created, organized and even digitized but still it cannot be considered available anywhere, anytime, for any citizen, business or other organization. Making governmental knowledge available to its beneficiaries requires the design, development and deployment of a Knowledge Registry, as the platform to cater for the formal description, composition and publishing of the governmental services canvas. Touching upon all knowledge management processes from knowledge capture, knowledge sharing, to knowledge creation, Chapter XIII goes beyond the methodology and tools used for developing such a system for the Greek Government, to the integration and the diffusion of e-Government knowledge with the help
of formal ontology definitions - capturing the core elements of the domain together with their main relationships.

Chapter XIV presents data mining, as a planning and decision support tool for biomass resources management to produce bio-energy. Furthermore, the chapter has tried to define the decision making problem for bio-energy production. The Decision Support System, that utilizes a data mining technique, e.g. clustering, integrated with other group of techniques and tools, such as Genetic Algorithms, Life Cycle Assessment, Geographical Information System, etc, is presented. Finally, a case study has been included that shows how to tackle the decision making problem using this technique.

The objective of Chapter XV is to provide an example of a user-friendly interface for knowledge management and information retrieval, through the use of virtual assistants in E-government applications. The chapter has also provided a short state of the art on-line virtual assistants technology, highlighting the knowledge management aspects. Two case studies from the Mexican state of Guanajuato, and the Federal Government Citizen’s Web Page are being presented and discussed. These case studies provide new insights into access methods, interfaces and ways to query and present information in e-government applications.

There is no doubt that e-government application in public administration and its productive use of ICTs would improve the interface between respective governments and their citizens in both service deliveries and provisions of basic needs. However, it is recognized that while there are many benefits that have been obtained by implementing e-government, there are many sectors of society that are not part of this growing electronic culture. Perhaps, economics, lack of access to the Internet and other technologies, low literacy levels and often lack of interest or willingness to use the new technologies, contributes to a country’s disparities in e-government practices. This way a gap is being created among the digitally reached and un-reached communities. Chapter XVI seeks to address the digital divide associated with e-government, which can serve as impediment for application of ICT. As a case study, the chapter explores the various initiatives that have been undertaken by the Malaysian government to bridge the digital gap.

Digitization initiatives and knowledge management have become an integral part in the changing global information society, where knowledge management comprises a range of practices used by organizations to identify, create, represent, and distribute knowledge. However, digitization initiatives refer to either jointly or collaborative efforts to translate existing library holdings whether in print, graphical, audio or combination of all, into digital format commonly known as electronic resources. As a case study, Chapter XVII discusses the problems of digitization, challenges and future opportunities for East African university libraries with focus on collaborative efforts and strategies backed up with policies for investments in ICTs training and integration of ICTs into the core university activities for effective knowledge management (KM) and information dissemination. It is argued in the chapter that digitization of library information will add value to more effective university KM, information access and use in multidisciplinary fields including local content.

Finally, as a case of social and political implications of data mining, Chapter XVIII put forwards application of data mining techniques in public sector administration. In the public sector, data mining was used as a means to detect fraud and waste, but in time it emerges as a measuring tool and means for improving program performance. This chapter discusses on issues like, improving services or performance, helping customer relations management, managing human resources, and detecting fraud, waste and abuse in the public sector administration. It also discusses on analysis of scientific and research information, detection of criminal activities, and detection of terrorist activities in public administration before providing future research directions on predictive analysis and diversity of data mining application domains.
CONCLUSION

The growing interest of contemporary researchers in data mining is motivated by a common problem across multifaceted disciplines, as how does one store, access, model, and ultimately illustrate and understand very large data sets? Historically, different aspects of data mining have been addressed by diverse disciplines. But, as a truly interdisciplinary text on data mining with social and political implications focusing knowledge management in e-government, blending the contributions of information science, computer science, social science, knowledge management, ontology, and algorithms deserves further studies and researches (Hand, Mannila & Smyth, 2001; Han & Kamber, 2005).

As unique its name carries, the book incorporates methodical and pragmatic research outcomes in the aspect of data mining accommodating socio-political interests. Implementation of knowledge based e-government plan to reach common citizens through ICT mediated techniques, especially utilizing data mining modus operandi may be directed toward the key target audiences, as such the community people. But, they would include decision-makers and policy initiators; the general public; industry and service communities; scientific and technical communities; academics; government and non-governmental organizations; public interest advocacy groups; civil society advocacy groups, donor agencies and development partners and international financial institutions.

In addition to these, as researches revealed data mining is the process of automatic discovery of patterns, transformations, associations and anomalies in massive databases, and is a highly interdisciplinary field representing the confluence of multiple disciplines, such as database systems, data warehousing, machine learning, statistics, algorithms, data visualization, and high-performance computing (COMP 290, 2003). Utilizing data mining techniques with social and political implications to improve e-governance require methodical knowledge management. Furthermore, knowledge management is desirable to facilitate information exchange and transaction processing with citizens, as well as to enable inter-government knowledge sharing and integration.

However, the core content generators (data generators) are the appropriate knowledge systems that connect people to reach other despite barriers of time, geography, culture, literacy, and even ownership of a telephone or computer or Internet. Therefore, broader audiences of this book will widely vary from individuals, researchers, scientists, academics, politicians, students, librarians, journalists and development practitioners. Finally, this book will generate tremendous impetus in terms of knowledge-based research initiations, thus will have highly acceptable scholarly value and at the same time potentially contribute to this very specific sector of research.

REFERENCES


