Chapter III
E–Government Knowledge Management (KM) and Data Mining Challenges:
Past, Present, and Future

LuAnn Bean
Florida Institute of Technology, USA

Deborah S. Carstens
Florida Institute of Technology, USA

Judith Barlow
Florida Institute of Technology, USA

ABSTRACT

Powerful data mining models and applications in e-government settings have the potential to bring major benefits to a wide range of stakeholders. As these models evolve, structural transitions occur within e-government to which include an evolution of managerial practices through knowledge management (KM). Unfortunately these efforts are vulnerable to a number of critical human interaction and behavioral components. This chapter examines e-government challenges regarding the linkages between data mining and KM over time, discusses the organizational development of e-government applications, and details both general and specific social, ethical, legislative, and legal issues that impact effective implementations. A final focus of the chapter is the potential strategic benefits of a risk-based approach that can be used to improve the core synergy of KM and data mining operations in e-government operations.

INTRODUCTION

Data mining in both the private and public sectors has focused on its ability to reveal patterns, relationships, and connections between behaviors or variables. From an e-government standpoint, data mining has the potential to bring major cost and convenience benefits to all stakeholders, including citizens and businesses. While the proliferation of data mining applications can provide very powerful data management tools, organizations employing these initiatives have been vulnerable to misunderstanding implementation limitations and oversight issues. When examined more closely, the majority of these limitations are closely linked to miscalculations about the human components that interface with these systems rather than the technology capabilities of a computer network. These human interaction and behavioral issues include such vulnerabilities as: (1) a dependence on data pattern and relationship determinations by analysts, (2) varying competencies and deficient analytic skills of specialists interpreting the data for knowledge management (KM) purposes, (3) falsely identifying relationships between variables that do not necessarily equate with causation, (4) misunderstanding information awareness issues of data quality and integrity, (5) mission creep (or the use of data for purposes beyond its primary collection), (6) privacy of data, (7) inadequate security of data (based on the interoperability needs of several organizations and user groups and the complexity of the distribution), (8) ethical breaches of users and suppliers of information, and (9) inadequate attention to laws and insufficient policies (Seifert, 2004).

In order to examine e-government's challenges regarding the linkages between data mining and KM over time, this chapter focuses on the organizational development of e-government applications that have evolved and currently interface with both general and specific social, ethical, legislative, and legal issues. The chapter also concludes with a future perspective regarding a risk-based approach for e-governments – one that improves the core synergy of KM and data mining operations.

BACKGROUND AND EVOLUTION OF E-GOVERNMENT AND DATA MINING

The history of e-government computing architectures and applications has seen similar developments to those evolving in the corporate world. While large, expensive mainframe computers of the 1960s and 70s were programmed to process inventory control and materials management data, the majority of costs were simply focused on capturing data in a machine-readable format and getting it into the system, accounting for over 80% of software project costs in the period from the 1960s through the early 1990s (Hazzan, Impagliazzo, Lister & Schocken, 2005).

Emphasis on data integrity concerns about the proper handling of data containing dates reached a climax in the late 1990s, causing many organizations to invest in high-performance database systems to avoid potential catastrophic errors when the year 2000 rolled around. A side effect of the rush to avoid Y2K problems, whether real or imagined, caused firms to implement large, expensive database management systems that led to better data capture, data storage, data processing and data sharing capabilities (Robertson & Powell, 1999).

With this new found wealth of cheap, accurate, timely, accessible and processable data from post-2000 information systems, the costs of deploying data mining applications was reduced, leading to new methods of explicitly capturing "knowledge" or "intelligence" from the large (and growing) data stores (Schwaig, Kane & Storey, 2005). The result is that data mining methods can be applied not only to an organization's own data stores but also to data copied, purchased or stolen from external sources or other government agencies.

Related Content

A Single Pass Algorithm for Discovering Significant Intervals in Time-Series Data

Sagar Savlaand Sharma Chakravarthy (2007). *International Journal of Data Warehousing and Mining (pp. 28-44).*

www.irma-international.org/article/single-pass-algorithm-discovering-significant/1788

A Semi-Automatic Annotation Method of Effect Clue Words for Chinese Patents Based on Co-Training

Na Deng, Chunzhi Wang, Mingwu Zhang, Zhiwei Ye, Liang Xiao, Jingbai Tian, Desheng Liand Xu Chen (2018). *International Journal of Data Warehousing and Mining (pp. 1-19).*

www.irma-international.org/article/a-semi-automatic-annotation-method-of-effect-clue-words-for-chinese-patents-based-on-co-training/215003

Pattern Comparison in Data Mining: A Survey

Irene Ntoutsi, Nikos Pelekisand Yannis Theodoridis (2007). *Research and Trends in Data Mining Technologies and Applications (pp. 86-120).*

www.irma-international.org/chapter/pattern-comparison-data-mining/28422

Incremental Algorithm for Discovering Frequent Subsequences in Multiple Data Streams

Reem Al-Mullaand Zaher Al Aghbari (2013). *Developments in Data Extraction, Management, and Analysis (pp. 259-279).*

www.irma-international.org/chapter/incremental-algorithm-discovering-frequent-subsequences/70801

Estimating Semi-Parametric Missing Values with Iterative Imputation

Shichao Zhang (2010). *International Journal of Data Warehousing and Mining (pp. 1-10).*

www.irma-international.org/article/estimating-semi-parametric-missing-values/44955