

Chapter 10

Ethical and Legal Data Mining: Paradigms of Organizational Development Practitioners and Organizational Psychologists

Ben Tran

Alliant International University, USA

ABSTRACT

In the process of data mining, techniques used, and results gathered, became ethically or even legally questionable due to the concern of the violation of individuals' privacy when specific information is obtained, manipulated, and disseminated by other entities without an individual's knowledge or consent. In other words, it is not the concern of what data is being mined, but more so regarding who is the miner of these data. This chapter is based on a review of the existing studies, which shows that not enough attention has been paid to the study of the "miner," as well as the ethical and legal qualifications of the miner, from the perspectives of organizational development practitioners (OD) and organizational psychologists (OP) in approaching for certifying or licensing. During the review, several cases are being included to justify the certifying or licensing approach, thus upholding ethical data mining, especially in relation to the socio-economic development of a community, or county. This research concludes with a few practical recommendations for both data miner(s) and entities that are involved with data mining.

INTRODUCTION

The use of databases to harness the power of data is not new, for it has occurred for centuries, and continues to grow rapidly for three primary reasons. First, organizations are beginning to understand that databases have the power to sort through massive amounts of meaningless data

and turn it into understandable and useful information. Second, the increased uses of database technologies make it possible to eliminate existing paper-based systems and digitize data, which ultimately makes it significantly easier to manage and effectively utilize it. Third, the major factor that has influenced database use is the advancement of digital storage technology.

DOI: 10.4018/978-1-4666-4078-8.ch010

The aim of most data mining applications is to extract useful new information from large volumes of structured and unstructured data, originating from various sources. The extracted information may be used for prediction and planning purposes or to improve the quality of business. Most data mining applications inevitably process data or information of a personal nature. Like most statistical methodologies, data mining, by itself, is ethically and legally neutral. This is particularly so because the term data mining is a generic one referring to a wide range of procedures, involving diverse data sets, and carried out for numerous purposes. As such, data mining is an umbrella concept. Thus, it is the techniques and some results of data mining that become ethically or even legally questionable. The privacy of an individual can be violated when specific information is obtained, manipulated, and disseminated by other entities without an individual's knowledge or consent (Olson, 2007).

Due to the violation of individuals' privacy, ethical and legal concerns in data mining arise, and are clustered around three different sets of issues. First, those related to the validity and appropriateness (method/how) of the statistical method employed. Second, those related to the use of microdata pertaining to identifiable individuals or mesodata used to target vulnerable population sub-groups (reason/why). Third, those that give rise to possible disservices to the statistics profession or to statistical agencies due to public backlash over actual or perceived shortcomings of specific data mining applications or other statistical techniques used in counterterrorism work (result/who). In other words, it is not the concern of what data is being mined, but more so regarding who is the miner of these data. Here implies the intricate notion of the miner, not the data itself, which is neutral and this, involves issues, such as legal and ethical aspects of the data mining that deserves appropriate attention from the data miners.

A review of the existing studies show that not enough attention has been paid to the study

of the "miner," as well as the ethical and legal qualifications of the miner, from the perspectives of organizational development practitioners (OD) and organizational psychologists (OP). The approach utilized here is an interdisciplinary approach from the social sciences, the OD and the OP, on ethical and legal factors via the three views of data mining: the function-oriented view, the theory-oriented view, and the procedure/process-oriented view. This study has tried to provide solutions to resolve the existing gap in the field of data mining regarding ethical and legal concerns. This research approach concludes with practical recommendations for both data miner(s) and entities that are involved with data mining and data miner(s).

According to Dr. Hakikur Rahman and Dr. Isabel Ramos, in the call for chapters, the target audience of this book, especially this chapter, consists of scientists, researchers, and practitioners working in the field of data mining, data warehousing, and database management and information systems technologies. The target audience also includes academics, research institutes, and individuals who are interested in this field, and most important of all, the large number of members in the public sector including government agencies, ministries, education institutions, health service providers, financial service providers, social service providers and other types of government, commercial and not-for-profit agencies, not just for computer science professionals and practitioners, specifically software engineers and data miners. Therefore, this chapter applies not just to computer science professionals and practitioners, specifically software engineers and data miners. It applies to researchers, practitioners, academics and individuals as mentioned above belonging to multi-disciplinary platforms. As such, as an organizational development and an organizational psychology practitioner, the author encourages the target audience to be open-minded, open-ears, and open in the finding the objective. The knowledge, information, and lessons covered here are unique

27 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:

www.igi-global.com/chapter/ethical-legal-data-mining/76263

Related Content

An Information-Theoretic Framework for Process Structure and Data Mining

Gianluigi Greco, Antonella Guzzo and Luigi Pontieri (2007). *International Journal of Data Warehousing and Mining* (pp. 99-119).

www.irma-international.org/article/information-theoretic-framework-process-structure/1796

Optimizing ETL by a Two-Level Data Staging Method

Xiufeng Liu, Nadeem Iftikhar, Huan Huo and Per Sieverts Nielsen (2016). *International Journal of Data Warehousing and Mining* (pp. 32-50).

www.irma-international.org/article/optimizing-etl-by-a-two-level-data-staging-method/168485

Efficient Identification of Similar XML Fragments Based on Tree Edit Distance

Hongzhi Wang, Jianzhong Li and Fei Li (2012). *XML Data Mining: Models, Methods, and Applications* (pp. 78-97).

www.irma-international.org/chapter/efficient-identification-similar-xml-fragments/60905

Statistical Sampling to Instantiate Materialized View Selection Problems in Data Warehouses

Mesbah U. Ahmed, Vikas Agrawal, Udayan Nandkeolyar and P. S. Sundararaghavan (2007). *International Journal of Data Warehousing and Mining* (pp. 1-28).

www.irma-international.org/article/statistical-sampling-instantiate-materialized-view/1776

GeoCache: A Cache for GML Geographical Data

Lionel Savary, Georges Gardarin and Karine Zeitouni (2007). *International Journal of Data Warehousing and Mining* (pp. 67-88).

www.irma-international.org/article/geocache-cache-gml-geographical-data/1779