Chapter XVII

Gaining Strategic Advantage Through Bibliomining: Data Mining for Management Decisions in Corporate, Special, Digital, and Traditional Libraries

Scott Nicholson
Syracuse University, USA

Jeffrey Stanton
Syracuse University, USA

ABSTRACT

Library and information services in corporations, schools, universities and communities capture information about their users, circulation history, resources in the collection and search patterns (Koenig, 1985). Unfortunately, few libraries have taken advantage of these data as a way to improve customer service, manage acquisition budgets or influence strategic decision making about uses of information in their organizations. In this chapter, we present a global view of the data generated in libraries, and the

variety of decisions that those data can inform. We describe 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. The chapter examines data sources and possible applications of data mining techniques in the library.

INTRODUCTION

For several decades, library and information services in corporations, schools, universities and communities have had the ability to capture information about users, circulation history, resources in the collection, and search patterns (Koenig, 1985). Collectively, these data can provide library managers more information about common patterns of user behavior to aid in decision-making processes. Unfortunately, few libraries have taken advantage of these data as a way to improve customer service, manage acquisition budgets or influence strategic decision making about uses of information in their organizations. The application of advanced statistical and data mining techniques to these kinds of data may provide useful ways of supporting decision making at any library where user, cataloging, searching and circulation interfaces are automated.

Use of data mining to examine library data records might be aptly termed bibliomining. With widespread adoption of computerized catalogs and search facilities over the past quarter century, library and information scientists have often used bibliometric methods (e.g., the discovery of patterns in authorship and citation within a field) to explore patterns in bibliographic information. During the same period, various researchers have developed and tested data mining techniques — advanced statistical and visualization methods to locate nontrivial patterns in large data sets. Bibliomining refers to the use of these techniques to plumb the enormous quantities of data generated by the typical automated library.

Forward-thinking authors in the field of library science began to explore sophisticated uses of library data some years before the concept of data mining was popularized. Nutter (1987) explored library data sources to support decision making, but lamented that "the ability to collect, organize, and manipulate data far outstrips the ability to interpret and to apply them" (p. 143). Johnston and Weckert (1990) developed a data-driven expert system to help select library materials, and Vizine-Goetz, Weibel and Oskins (1990) developed a system for automated cataloging based on book titles (also, see Morris, 1991; Aluri & Riggs, 1990). A special section of Library Administration and Management ("Mining your automated system") included articles on extracting data to support system management decisions (Mancini, 1996), extracting frequencies to assist in collection decision making (Atkins, 1996) and examining transaction logs to support collection management (Peters, 1996).

More recently, Banerjeree (1998) focused on describing how data mining works and ways of using it to provide better access to the collection. Guenther (2000) discussed data sources and bibliomining applications, but focused on the problems with heterogeneous data formats. Doszkocs (n.d.) discussed the potential for applying neural networks to library data to uncover possible associations between documents, indexing terms, classification codes and queries. Liddy (2000) combined natural language processing

14 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/gaining-strategic-advantage-through-bibliomining/27920

Related Content

Clubhouse Experience: Sentiment Analysis of an Alternative Platform From the Eyes of Classic Social Media Users

Ipek Deveci Kocakoçand Pnar Özkan (2022). *Data Mining Approaches for Big Data and Sentiment Analysis in Social Media (pp. 244-264).*

www.irma-international.org/chapter/clubhouse-experience/293159

Machine Fault Diagnosis and Prognosis using Self-Organizing Map

Kesheng Wang, Zhenyou Zhangand Yi Wang (2014). *Biologically-Inspired Techniques for Knowledge Discovery and Data Mining (pp. 129-148).*

www.irma-international.org/chapter/machine-fault-diagnosis-and-prognosis-using-self-organizing-map/110457

A New Approach for Supervised Dimensionality Reduction

Yinglei Song, Yongzhong Liand Junfeng Qu (2018). *International Journal of Data Warehousing and Mining (pp. 20-37).*

www.irma-international.org/article/a-new-approach-for-supervised-dimensionality-reduction/215004

Approximate Range Querying over Sliding Windows

Francesco Buccafurri (2009). *Handbook of Research on Text and Web Mining Technologies (pp. 273-287).*

www.irma-international.org/chapter/approximate-range-querying-over-sliding/21729

A Hyper-Heuristic for Descriptive Rule Induction

Tho Hoan Phamand Tu Bao Ho (2007). *International Journal of Data Warehousing and Mining (pp. 54-66).*

www.irma-international.org/article/hyper-heuristic-descriptive-rule-induction/1778