emeraldinsight

Library Hi Tech

Library management and innovation in the Big Data Era
Shan Liu, Xiao-Liang Shen,

Article information:

Users who downloaded this article also downloaded:

For Authors

If you would like to write for this, or any other Emerald publication, then please use our Emerald
for Authors service information about how to choose which publication to write for and submission
guidelines are available for all. Please visit www.emeraldinsight.com/authors for more information.

About Emerald www.emeraldinsight.com

Emerald is a global publisher linking research and practice to the benefit of society. The company
manages a portfolio of more than 290 journals and over 2,350 books and book series volumes, as
well as providing an extensive range of online products and additional customer resources and
services.

Emerald is both COUNTER 4 and TRANSFER compliant. The organization is a partner of the
Committee on Publication Ethics (COPE) and also works with Portico and the LOCKSS initiative for
digital archive preservation.

Guest editorial

Library management and innovation in the Big Data Era
Introduction
Nowadays, big data has become a significant issue across different areas and received considerable attention in both research and practice. Although scholars in library and information science (LIS) discipline have claimed that big data presents a huge opportunity for library research (Gordon-Murnane, 2012), the role of big data in facilitating library management and innovation remains unclear and needs further investigation. As witnessed, big datadriven library innovations provide personalized, remote, real-time, and virtualized services to the library users. In addition, the rapid increase in volume, veracity, velocity, and variety of library data generated by different library tools offers innovative ways of understanding interactions with users in the library environment (Nicholson and Bennett, 2016).

This special issue aims to address not only library management and innovation issues exposed by library innovative applications, technology, and services, but also technical and managerial approaches, methodologies, and solutions that would overcome the challenges encountered by the librarians in the era of big data. This special issue makes some significant and original contributions to both library management and LIS community in general. For library practitioners and researchers, this special issue utilizes big data thinking and data analytics approach to address the underlying management problems and promote various innovations in the library. For LIS scholars, this special issue is an early attempt to understand how library data can be acquired, preserved, processed, and applied to generate considerable value and true insights.

A framework of big data-driven library management and innovation
Big data enables library to be smart and user-friendly by providing personalized and intelligent services. Generally speaking, library big data can be categorized into two groups: catalogue and process/transactional data. Catalogue data mean the inherent data and information of library files, while process data are often generated through the process of library management and service or created by library users. The former group of data generally contains documental, bibliographical, and funding data, while the latter group includes log, user, and record data. The analytics of library big data support tremendous digital library innovations, such as personalized recommendation services and library user behavior/habit analysis, which generate substantial value and insights for librarian, user, and services. Values for librarian are manifested in the changes and benefits provided by intelligent big data analytical techniques to librarians and digital management processes that enables library to provide competitive products and services with minimal costs. User value is embedded in the improvement of library user experience and satisfaction of users. Service value includes improving service and process quality and efficiency with the analysis of library big data in its varying forms.

Digital library innovation also requires library to provide a robust and intelligent digital management system. Due to the high volume, variety, velocity, and veracity of big data, the development of library management system requires not only the design of a new architecture, but also the application of digital technology in managing library big data, such as data acquirement, preservation, and processing. Therefore, both digital library management and digital library innovation constitute a transition closed-loop system, where library innovation drives library management, which in turn provides management

systems and necessary support for library innovation. Based on the above-mentioned discussions, we develop a framework of big data-based library management and innovation (see Figure 1). This framework can be applied in the future research and practice regarding library management and innovation in the big data era.

Summary of contributions to special issue

A total of 26 submissions were received and ten articles were accepted in this special issue. Each article is related to library big data and contribute to the library management and innovation literature by addressing certain aspects of the framework. In particular, four articles have employed catalogue data to investigate digital library innovation. Wu, Cai, Jin and Dong (2018) have collected scientific funding data of 193,517 projects of the National Natural Science Foundation of China, which are often overlooked by the librarians. They used a co-occurrence network analysis to develop an interdisciplinary knowledge flow network. They identified heterogeneous knowledge distribution across disciplines, but each discipline has mutual influences on each other. Through such big data analysis, libraries can provide funding consultation service to help funding application. Fang *et al.* (2018) focused on bibliographical data and used a probabilistic generative topic model (i.e. latent Dirichlet allocation) to analyze hot and cold topics from library electronic references. The topics identified by this automatic text analysis are accordant with the results discovered by human experts, thus generating a great value for libraries to provide advanced services for users. Zhu *et al.* (2018) focused on unstructured publication data of library information science journals. They used ensemble empirical mode decomposition, artificial neural networks, and theme logic structure methods to identify preferences of editors, reviewers, and authors. Their finding enables journal mangers and potential authors to predict the research characteristic tendencies of journals. Chang (2018) explored documental data and integrated genealogical data with migration analysis to examine the migration patterns. They also designed a genealogical visualized research platform to display the paths of migration and family history. The combination of big data and geographic information system provides insights for the library to provide an innovative genealogy service.

Process data can also be utilized and analyzed for digital library innovation. Yi *et al.* (2018) employed book-borrowing record data and used an association rule-mining algorithm to provide personalized recommendation service for the readers. This service is based on readers' historical borrowing records and delivers a great value for library users to read their favorite books. Liu *et al.* (2018) employed a data set from CiteULike and developed a collaborative matrix factorization mechanism for group recommendation under Hadoop framework. Their proposed methods enabled the interactions between groups
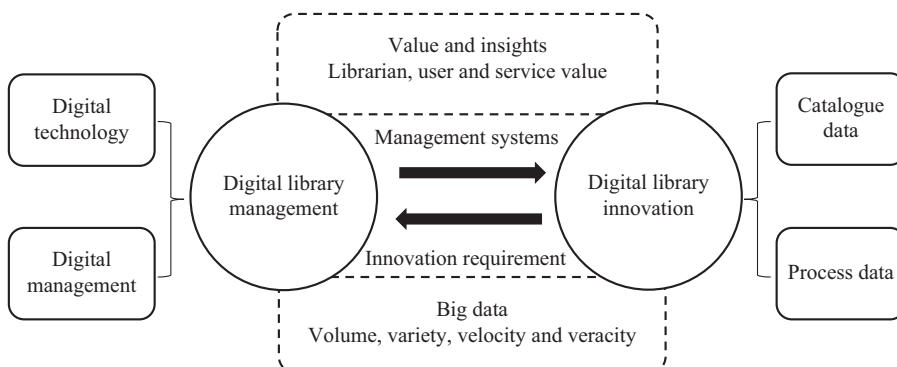

Figure 1.
Framework of library management and innovation in the big data era

and members and improved group activity effectiveness and service quality of online library systems. This method is also beneficial to scientific collaboration and membership retention. Wu, Liang and Bi (2018) identified query reformulation patterns to understand users' cross-device research behavior by analyzing online public access catalogue data. They discovered the best pattern and time interval of device transition and identified the reason that triggers the transition. Their research added new value to the cross-device search and interaction, which enables library to provide better digital services and devices.

Digital library innovation is closely and inseparably related to effective management of library process and big data so that library can be highly smart. Simović (2018) proposed a big data smart library that can increase continuous learning process through recommendations and improve user service. He also designed a system for collecting, analyzing, processing, and visualizing big data from various sources. This smart library system not only creates new value to assist users in finding personalized content, but also enhances library operations by integrating big data technology. Gerrard *et al.* (2018) focused on digital preservation for big data and investigated how such system architectures support business analysis of large-scale data. They proposed a new architecture in preserving big data and improving the usefulness and ability of library to process digital resources at big data volume. Bhat (2018) explored the prospect of current storage technologies for long-term preservation of big data in digital libraries and proposed emerging preservation technologies as a long-term viable solution. This article provided insights to frame sustainable policies and understand the technical details to select data preservation technologies.

Conclusions

This special issue investigates how libraries manage and harness big data to boost innovations and create values for librarian, user, and service. We present a general framework of library management and innovation in the big data era in dealing with two kinds of library big data, namely, catalogue and process data. The articles published in the special issue highlight the analysis of catalogue and process data in delivering innovative digital library services for users. It also provides system architecture design and smart digital solution for innovative library management. Future research can continue to focus on big data-driven digital library management and innovation to gain value and insights through data analytics.

Shan Liu
School of Management, Xi'an Jiaotong University, Xi'an, China, and
Xiao-Liang Shen
Economics and Management School, Wuhan University, Wuhan, China

References

Bhat, W.A. (2018), "Long-term preservation of big data: prospects of current storage technologies in digital libraries", *Library Hi Tech*, Vol. 36 No. 3, pp. 539-555.

Chang, C.C. (2018), "Hakka genealogical migration analysis enhancement using big data on library services", *Library Hi Tech*, Vol. 36 No. 3, pp. 426-442.

Fang, D., Yang, H., Gao, B. and Li, X. (2018), "Discovering research topics from library electronic references using latent Dirichlet allocation", *Library Hi Tech*, Vol. 36 No. 3, pp. 400-410.

Gerrard, D.M., Mooney, J.E. and Thompson, D. (2018), "Digital preservation at big data scales: proposing a step-change in preservation system architectures", *Library Hi Tech*, Vol. 36 No. 3, pp. 524-538.

Gordon-Murnane, L. (2012), "Big data: a big opportunity for librarians", *Online*, Vol. 36 No. 5, pp. 30-34.

Liu, Y., Yang, L., Sun, J., Jiang, Y. and Wang, J. (2018), "Collaborative matrix factorization mechanism for group recommendation in big data-based library systems", *Library Hi Tech*, Vol. 36 No. 3, pp. 458-481.

Nicholson, S.W. and Bennett, T.B. (2016), "Dissemination and discovery of diverse data: do libraries promote their unique research data collections?", *International Information & Library Review*, Vol. 48 No. 2, pp. 85-93.

Simović, A. (2018), "A Big Data smart library recommender system for an educational institution", *Library Hi Tech*, Vol. 36 No. 3, pp. 498-523.

Wu, D., Liang, S. and Bi, R. (2018), "Characterizing queries in cross-device OPAC search: a large-scale log study", *Library Hi Tech*, Vol. 36 No. 3, pp. 482-497.

Wu, J., Cai, J., Jin, M. and Dong, K. (2018), "Embedding funding consultation in library services: a co-occurrences network analysis of knowledge flow in scientific funding", *Library Hi Tech*, Vol. 36 No. 3, pp. 378-399.

Yi, K., Chen, T. and Cong, G. (2018), "Library personalized recommendation service method based on improved association rules", *Library Hi Tech*, Vol. 36 No. 3, pp. 443-457.

Zhu, Q., Wu, Y., Li, Y., Han, J. and Zhou, X. (2018), "Text mining-based theme logic structure identification: application in library journals", *Library Hi Tech*, Vol. 36 No. 3, pp. 411-425.