Big Data: Opportunities and Challenges in Libraries, a Systematic Literature Review

Emmanouel Garoufallou and Panorea Gaitanou

Currently "Big Data" is an emerging field that presents several Information Technology challenges regarding the capture, storage search, structure, and visualization of this data. The real challenge for organizations is to find ways to extract value from it and provide better services to their clients. The data generated in academic and other institutions is vast and complex. Libraries face new challenges as they seek to determine their role in the handling of Big Data within their organization and use it to develop services. Thus, in most organizations, libraries will not have the knowledge to build new services unaided. Furthermore, libraries have always been information handlers and technology adopters; therefore, Big Data technologies will certainly affect their context. The purpose of this paper is to explore all these issues through a systematic literature review, unveiling the theories that underpin the paper's argument. It attempts to answer several research questions, such as how librarians are involved in the Big Data era? And what are the future research developments of Big Data within the library context? The study considered only papers published between 2012 and 2018 in English and presents the collected literature by grouping them according to the type of library each paper refers to. Thus, it identifies new and evolving roles in the context of all types of libraries. In addition, the study presents several interesting tables, which aim to help librarians locate relevant articles that will inform their practice and guide service development for users of large and complex datasets.

Introduction

Currently, many information professionals are considering the library's involvement with Big Data. What is the library's role with any data and information? The answer is simple: to identify and select valuable resources; organize, describe, and preserve them; and finally provide access to their patrons. While dealing with Big Data, librarians are asked to involve themselves at earlier stages of the information cycle, while traditionally they focused on handling the products of research that was filtered through publication.

Since 2012, references to the term "Big Data" have become more frequent in the headlines of newspapers, proprietary magazines, and academic journals in many disciplines.[1] Big Data is valuable to most subject areas; its capabilities, uses, and applications are varied[2] and can

*Emmanouel Garoufallou is Associate Professor, Department of Library Science, Archives and Information Systems, School of Social Sciences at the International Hellenic University, Greece; email: garoufallou@gmail.com. Panorea Gaitanou, Alcalá University, Spain; email: rgaitanou@gmail.com.

yield surprising insights.[3] In the new digital era, in the fourth industrial revolution, Big Data is everywhere. Chung and Kim, by presenting the results from a survey of the World Economic Forum (WEF), concluded that Big Data, together with the Artificial Intelligence, the Internet, and IoT, as well as "Blockchain, 3D Printing, Sharing Economy and biotechnology," were among the main fields that will significantly affect our lives upon fourth industrial revolution.[4] Therefore, it is important to prepare for the changes that will soon or later occur in society. So, where do libraries stand with these changes? Many questions arise on how libraries will follow, implement, integrate, and work side-by-side with these new innovations and challenges. In this paper, one particular question that arises is: to what extent is Big Data the next major challenge for library and information science?[5] The main goal of data science and information professionals is to transform large and messy datasets into actionable knowledge through analytical thinking procedures. Libraries should therefore aim to facilitate knowledge creation in their communities.[6]

This paper explores the ways Big Data affects libraries and how librarians are prepared for supporting data in their services. Big Data is certainly a new area of academic research; thus, it is essential to explore the aspects that impact libraries and the new challenges it presents. The field of networks and digital technologies is undoubtedly dynamic and rapidly developing and in turn this has led to the continuously increasing volumes of information. Within the Big Data context, the traditional library service concept has changed, as effective library collection development requires the effective analysis of the needs of library patrons.[7] In addition, with the implementation of Big Data technologies, new knowledge can be gained and new services may be provided, adding value to existing ones.[8] Furthermore, librarians can employ big data analytics to evaluate and improve library services[9] and provide more high-quality, targeted services[10] characterized by a self-adaptive personalized information system and a knowledge information service for auxiliary decision-making.[11]

This paper reviews the literature on Big Data and libraries, identifying articles focused on how Big Data can help libraries to better understand their patrons, by implementing Big Data technologies and thereby adding value. The review of the existing evidence on the implementation of Big Data in this context and all the ideas and thoughts about future plans and directions will inform the design of more user-centered tools and strategies while highlighting some of the main challenges and problems related to the application of Big Data.

The rest of the paper is structured as follows: Section two presents an overview of what Big Data represents for libraries, its potentials and problems. Section three outlines the methodological approach undertaken for conducting this systematic review, while Section four includes the results of this study respectively. Finally, conclusions are presented in the last section.

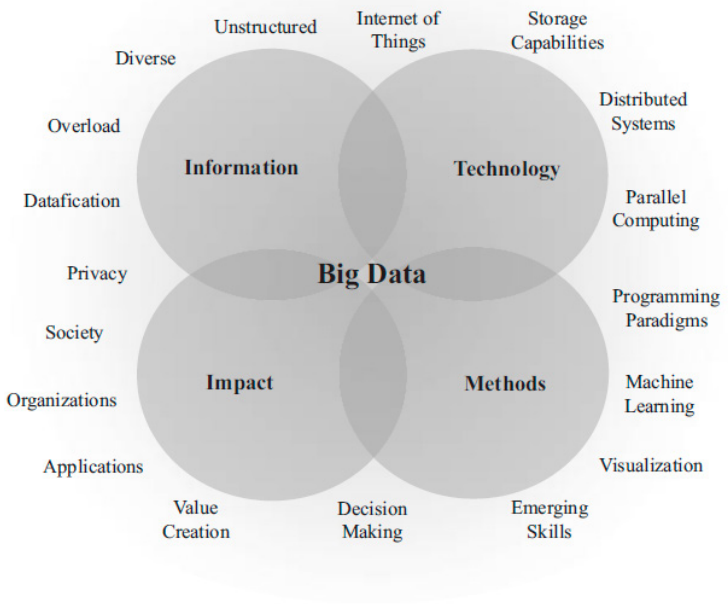
Big Data Overview: Definitions, Potentials, and Problems

There are no clear definitions for the term "Big Data." The term has been around since the 1990s, with some giving credit to the work of Mashey.[12] In 2001, Laney described the characteristics of Big Data as data that cannot be processed by traditional data management tools.[13] Dumbill offered the following conceptual definition:[14] Big Data "is data that exceeds the processing capacity of conventional database systems; the data is too big, moves too fast, or doesn't fit the structures of your database architectures; to gain value from this, you must choose an alternative way to process it." In general, Big Data can be defined as data being generated constantly, automatically and rapidly[15] and is a much more complex issue than just massive amounts of data.

The modern term Big Data refers to "data whose scale, diversity and complexity require new architectures, techniques, algorithms, and analytics to manage it and extract value and hidden meaning from it."[16] In particular, Big Data is a "combination of four very important characteristics: volume, velocity, variety and veracity":[17] volume refers to the amount of data; velocity refers to data in motion and more specifically to the speed at which data is created, processed and analyzed; variety is about managing the complexity and heterogeneity of multiple datasets, including structured, semistructured, and unstructured data; finally, veracity refers to data uncertainty and to the level of reliability/quality associated with certain types of data. De Mauro et al. present a very interesting attempt to define the term Big Data,[18] as shown in figure 1. Klapwijk classifies Big Data into three classes of "datafication":[19] a) directed data, like surveillance data; b) automated data, like device-generated data; and finally, c) volunteered data, such as social networks data. Big Data sets can contain unstructured data, such as email messages, photographs, or postings on internet forums.[20]

The rapid speed at which this data is being created adds to its dynamic capabilities,[21] although the actual value of raw data is still unknown. Big Data analytics have been applied in several industries (for example, the energy and medicine domains). Particularly, as far as the healthcare domain is concerned, Gaitanou et al. aimed to present a systematic literature review to determine the extent to which Big Data applications in healthcare systems have managed to improve patient experiences and clinicians' behavior, as well as the quality of care provided to patients.[22] Several other surveys[23] indicate that, in the near future, Big Data systems and technologies will help doctors to provide therapies based on diagnoses given by these systems, by helping to unlock key insights in the move to value-based care and precision medicine. Furthermore, there is an extensive discussion about Big Data and libraries as part of the Special Track on Digital Libraries, Information Retrieval, Big, Linked, Social & Open Data as part of the MTSR Conference (Metadata and

FIGURE 1
Definition of Big Data According to De Mauro et al.



Semantics Research Conference).[24] Nevertheless, the authors should mention that there are also several problems that need to be addressed in the Big Data revolution. There are several legal and policy issues.[25]

The main issues are privacy and data ownership:[26] for instance, in healthcare, social networking and government domains contain large amounts of sensitive information. Moreover, Big Data can be misleading[27] or meaningless. Bad data-based decisions can make data not approachable and valuable, thus leading decision makers down the wrong path.[28] Also, there are infrastructure issues that need to be addressed while managing Big Data.[29] Large numbers of datasets will be created: big storage spaces will be needed and network capacity will be increased. Finally, the complexity of the data must be seriously considered.[30] Lemieux et al. identified several other big data challenges with visual analytics that refer to: unavailability, fragmentation of data, poor data quality, and lack of data standardization.[31]

Methodology

This study sought to identify and review the literature on Big Data and its role within the library context. The review has followed the rules of a systematic review[32] by using the following research strategy: <library-related keywords> AND <Big Data–related keywords>. Library-related keywords and Big Data–related keywords were combinations of the following: [library OR libraries OR "library world" OR "library context" OR "information centers" OR librarians OR "information scientists"] AND ["Big Data" OR "massive data" OR "big datasets" OR "big datascales" OR "complex datasets" OR "large datasets" OR "large amount of data" OR "raw data"]. Queries were first submitted to Scopus, the largest abstract and citation database of peer-reviewed literature that includes scientific journals (such as Digital Library Perspectives, Electronic Library, International Information and Library Review and others), books and conference proceedings that give access to premium quality research information from 24,600 active titles and 5,000 publishers around the world; then to Google Scholar and to LISTA database, to expand the research results and include all relevant publications that were not found in Scopus. The research for identifying relevant literature was conducted between August 2018 and October 2019. The analysis covered the period 2012–2018. Only articles written in English were included. Finally, the authors also reviewed the reference lists of the included papers.

Inclusion and Exclusion Criteria

The literature review included all types of publications that focused on the role of Big Data within the library context: conference papers, journal articles, presentations, book chapters, and bulletin articles. Moreover, it included opinion, position, or concept papers on the use of Big Data in libraries, as well as review papers, and papers reporting on planned, but not actually implemented, studies. Finally, articles before the year 2012 were excluded. Also, four papers[33] published in the beginning of 2019 were not included in this research. This indicates that new research on this field emerges continuously.

Data Extraction

Following a review of titles and abstracts, the search identified 148 potentially relevant articles. Once all relevant papers were identified, a thorough reading was required to as-

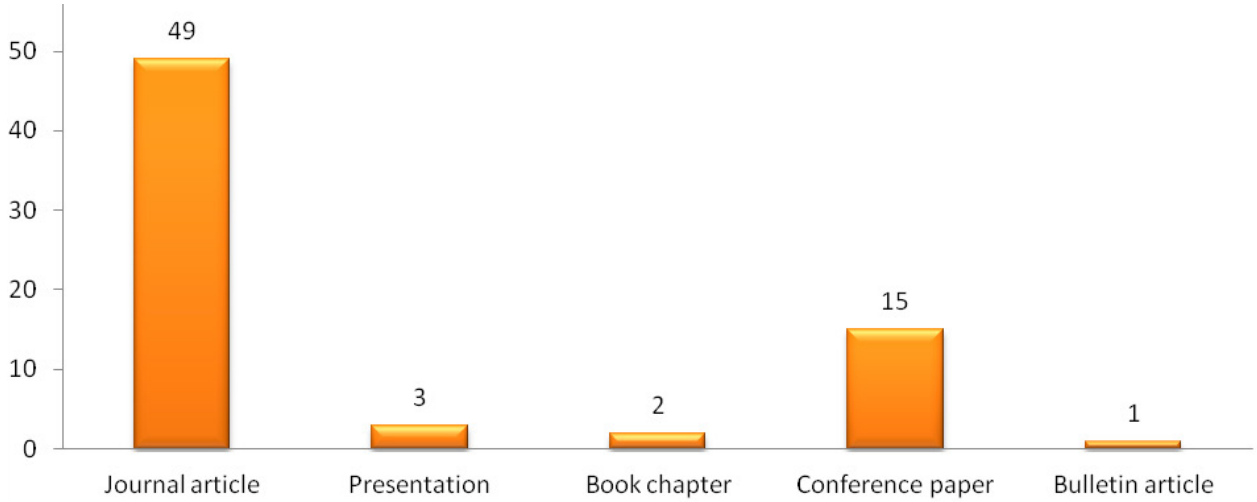
sess each one's importance, relevance, and quality. Each paper's aim and objectives were highlighted, and a list with all publication dates, titles, and aims was created. Finally, 70 papers satisfied the inclusion criteria. As shown in appendix A (see table 1), for each article included in the study a record was kept in an Excel spreadsheet with the following information:

- The article (reference)
- Year of publication of the article
- Type of paper (a variety of materials was identified such as journal articles and proceedings papers)
- Themes of the papers (keywords)
- Type of Big Data analyzed or planned to be analyzed (such as research data, log data, library statistical data)
- Type of library it refers to (such as academic libraries and public libraries)

From the analysis of the data extracted above, we generated selected infographics based on appendix B (see tables 2, 3, and 4) to illustrate the following:

An initial categorization can be provided according to the type of publication (see figure 2). In particular, a variety of materials was identified, such as journal articles, conference papers, presentations, book chapters, and bulletin articles. More specifically, the majority of the identified papers (49 in total) were journal articles, 15 were conference papers, three in the type of presentation, two were book chapters, and the final one was a bulletin article.

FIGURE 2
Identified Papers' Type of Publication



Furthermore, as already mentioned, the searches for identifying relevant literature were conducted between August and October 2018. In this context, the identified relevant literature covered papers published between 2012 and 2018, as shown below (see figure 3).

Finally, categorization according to the type of the library it refers to was considered particularly important for highlighting the aim and objectives of this review (see figure 4). The presentation and discussion of the relevant literature are analytically presented in the following section.

FIGURE 3
Number of Identified Relevant Articles per Year

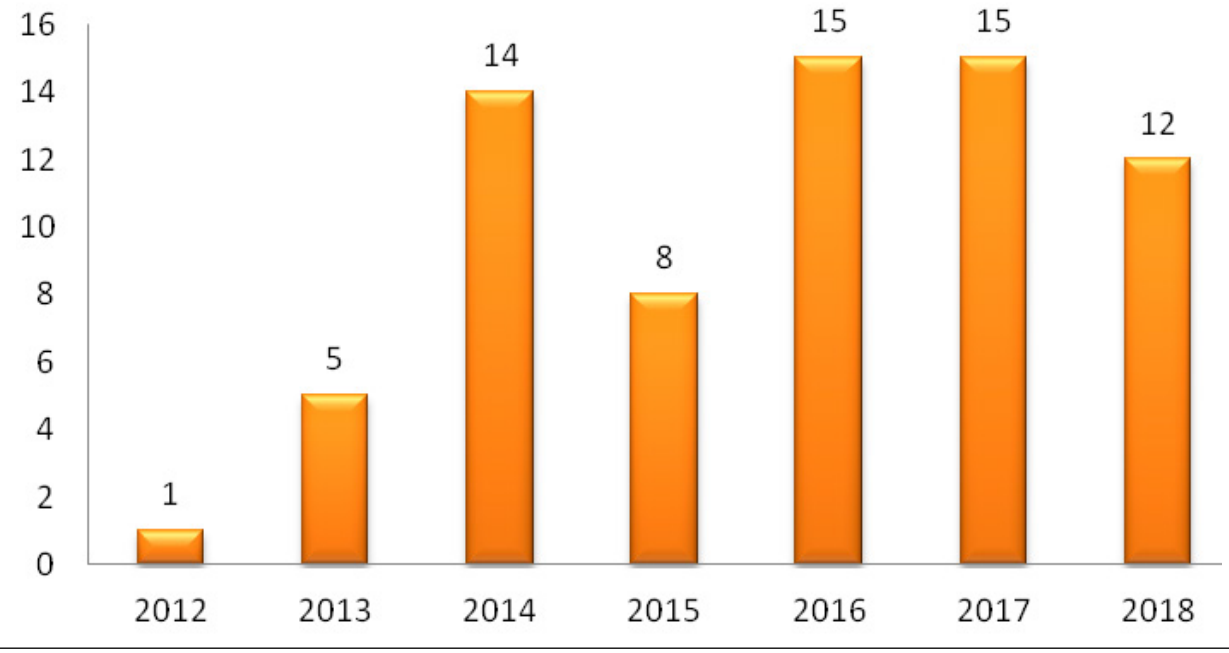
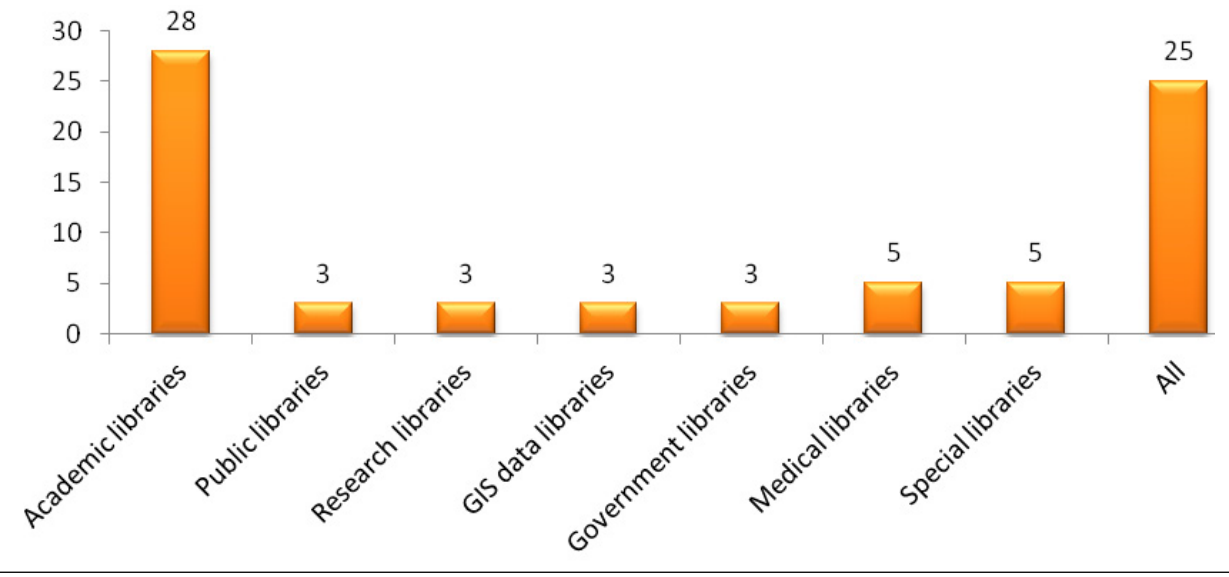


FIGURE 4
Library Type of Identified Relevant Papers



Results

As library's role is to identify, evaluate, select, organize, and describe resources, as well as preserve, disseminate, use, and reuse resources and information, library's role is getting more complex today than ever. This role for library, librarians, and information scientists as information providers finds them in a crossroad that forces them to choose if they will play a central role in the meta-fourth industrial revolution as central-information providers or if they will keep a profile as organizations that continue to provide "traditional" services to patrons. The dilemma is big, as big are the Big Data and the changes that will bring to our society.

Our research indicated several articles that present research and ideas regarding surveys or implementation of Big Data technologies within the library sphere. More specifically, we categorized these articles by highlighting the type of library to which they refer. Thus, the review showed that 25 articles refer to all types of libraries. Big Data management in libraries within organizations is very essential as librarians need to help their patrons to understand what Big Data is and how they can successfully manage them so as to achieve their goals. Zhan and Widén aimed to generate an overall understanding of Big Data limited to librarianship, because of its unique position in managing and using Big Data.[34] Hoy mentioned that libraries are involved through collection development and preservation of datasets for future users. As more users become interested in working with Big Data, they will need guidance and material to work with.[35] Kumar and Priyadarsini discussed the characteristics of datasets in libraries and conducted a review of research work on library Big Data, summarizing the applications in this field.[36] They noted that research data increase too fast as more and more researchers use huge collections to mine and organize information in novel ways. Thus, Big Data can certainly help libraries make more cost-effective, innovative decisions or recommendations that users wish to have. They mentioned that Big Data can help to improve the overall user experience, as well as user satisfaction with library service. Similarly, Xu et al. discussed the characteristics of datasets in libraries and argued against the popular notion that there is insufficient research on libraries and Big Data. In addition, the status of Big Data research in library in China was explored.[37] The authors supported that library Big Data could also more effectively serve researchers or ordinary users better. They also observed that librarians still are not sure how to integrate Big Data into library data, and they needed to possess a deep understanding of how to transform, analyze, and present data to facilitate knowledge creation. Harper and Oltmann examined how Big Data is applied in the corporate world, highlighting Big Data's usage as it applies to the library system.[38] Leetaru's speech offered librarians a glimpse of various Big Data applications using images, maps, and tweets.[39]

Kalpwijk believed that libraries have to start thinking data-analytically. Value can be extracted from small data by scaling them up into larger datasets, for reuse through digital data infrastructures.[40] Wang et al. supported the idea that librarians should know how to make big datasets more useful, visible, and accessible, and, via Big Data technologies, researchers can look at data in new ways.[41] Li et al. provided a comprehensive overview leading to a Big Data application framework in libraries.[42] Ahmed and Ameen presented a scientometric analysis of papers indexed in Thomson Reuters' ISI Web of Knowledge, using Vosviewer software to explore the research trends associated with Big Data in the field of library and information science.[43] Cervone provided a practical example of how Big Data and analytics can help libraries and information organizations to evaluate the effectiveness of social media efforts to improve the perception of products and services in information organizations.[44]

Qin proposed five stages in the process of getting Big Data information: preparation, preprocessing, data mining, output, and evaluation, among which data mining is the core of Big Data thinking. In this context, the model of library information management system is constructed.[45] Furthermore, the organization management system design process of the personalized information service system at a big library data environment is introduced in detail by Hao.[46] Bhat explored the prospects of current storage technologies for long-term preservation of Big Data in digital libraries and concluded that current technologies are not

viable for long-term preservation and cannot fulfill all the storage demands or alleviate the financial expenditures of digital libraries.[47]

Golub and Hansson illustrated LIS developments and challenges particular to the world of Big Data to readers from computer science and related disciplines.[48] Long believed that in the future the digital library will replace the traditional library. Therefore, it is obvious that services provided in the era of Big Data will become a new hot spot for scholars.[49] Olendorf and Wang explored the benefits, costs, and risks associated with the use of Big Data. They also provided use cases, guidance to getting started, and a brief outline of tools and resources for libraries to work with Big Data.[50]

Furthermore, Sun and Ma discussed the features of libraries in the Big Data era and its influence on information security of libraries, and they proposed Big Data-related scientific problems that should be solved in terms of information security.[51] Wei et al. reported that Big Data technologies can identify hidden values behind the data through systematic organization and analysis of massive structured, unstructured, and semistructured data to predict the future development of libraries and promote their better development.[52] He and Zhang used Big Data technologies to push information to readers to improve the quality of library services.[53] Ping referred to Big Data mining and association rules–based personalized recommendation technology and supported two approaches that can increase the lending rate of collections.[54] Campbell and Cowan addressed the concerns of protecting library users' privacy and suggested that linked data technologies, with their ability to lead searchers through self-directed, open inquiry, are superior to Big Data technologies in the navigation of the paradox between openness and secrecy.[55] In addition, regarding the Digital Humanities field, Poole and Garwood referred to the DID3 projects as a case study to explore the roles that librarians and archivists play in data-intensive, interdisciplinary, and international DH projects.[56]

Libraries started dealing with Big Data when they automated their cataloguing functions and circulation systems through OCLC and similar data-sharing organizations. Thus, one of the first examples of Big Data in the bibliographic and library world could be WorldCat, the world's largest network of library content and services, which has the potential to support knowledge discovery in many areas. Teets and Goldner focused on WorldCat and gave an explanation of how OCLC is actively involved with Schema.org to make this data used throughout the web.[57] Finally, Green and Panzer explored how WorldCat data can inform the development of the DDC (classification analytics) and how DDC-classified content in WorldCat can shed light on the bibliographic world itself.[58]

The rest of the identified papers can be categorized as below:

Big Data in Academic Libraries

Academic libraries have a long history of collecting data, reporting their analyses, and compiling them into library statistics as a way to assess the library's resources and performance.[59] Recently, the rise of Big Data has made several data collection tasks easier and faster and has engaged libraries in complex data analysis.[60] A number of papers (28 in all) refer to academic libraries. Deng analyzed the characteristics of Big Data and its influence to the university library context.[61] Under the background of the Big Data era, Wang believed that university libraries should collect and analyze more native data and provided more tools for knowledge discovery to provide qualified information service and meet the need of users.[62] Luo et al. explored cloud services and Big Data processing technologies within the university library context,

paying special attention to issues such as privacy protection and standardization.[63] Gerrard et al. focused on digitized images and research data and explored the future relationship between Big Data and digital preservation that has been guided by hands-on research activity.[64] Lu et al. presented a Big Data age-featured innovative health information services model in university libraries.[65] Yi et al. discussed Big Data and the relevant theoretical knowledge of Big Data mining, by designing the personalized recommendation system service model of the university library based on clustering analysis and association rules theory.[66] Sun et al. used Big Data technology to mine, identify, organize, and analyze the implied reader behavior in structured and semistructured data information to improve library services and resources to achieve the optimum configuration.[67] Allalouf used Big Data technologies to learn about searching trends in library catalogues and thus improve library resources.[68]

Chen et al. proposed a new search behavior assessment methodology by introducing analytics techniques in library transaction logs.[69] The authors mentioned that log analysis has two serious disadvantages: a) the lack of contextual information, such as users' motivations (for example, book use patterns are probably influenced by changes to essay topics); and b) the fact that one can reach oversimplified conclusions (amateurization of "Big Data" analysis). Golub and Hansson believed that the mission of librarianship is to facilitate knowledge creation in communities; therefore, it is necessary for librarians to transform raw, messy data into actionable knowledge that can be used by decision makers.[70] Fister mentioned that as collecting and analyzing data is becoming part of the curriculum nowadays, a significant challenge is presented for academic libraries, which will likely require making some difficult decisions about staffing and resource allocation.[71] Jantti presented the case of the University of Wollongong Library in Australia, which, since 2010, has sought to integrate library usage data within enterprise reporting systems to better understand the value and impact for students using library resources.[72] Li analyzed the impact of Big Data on colleges and universities libraries,[73] while Goldberg et al. explored the existence of Big Data in a specific type of academic libraries, the map and Geographic Information System (GIS) data libraries (MGDLs), which can be analyzed separately. They highlighted the evolving landscapes within which MGDLs have to operate by examining how their roles and operational organization will be impacted.[74] Kong referred to challenges regarding geospatial data management and curation, which include the application of Big Data, the emergence of Web GIS, and the advancement of cyberinfrastructures.[75] Weessies and Dotson aimed to describe several data-rich GIS Lite tools available in the library market and explain how these products have managed or failed to meet the needs of several real-life college class situations.[76]

Chang presented an innovative library service, which enhances the Hakka genealogical migration analysis using Big Data.[77] In this way, the data connection, spatial distribution characteristics, and trends in genealogical information are explored. Halperin and Lusk illustrated the various stages of the Specific Context Benchmarking (SCB) protocol using the data produced by the Academic Research Libraries to enhance insights regarding the details of the operational benchmarking context and so offer generalizations needed to encourage adoption of SCB across other functional domains.[78] Tuppen et al. introduced the "Big Data History of Music" project, a collaboration between Royal Holloway, University of London, and the British Library, and described the challenges librarians face to access large music datasets.[79]

Linlin attempted to analyze the application of information processing technology in university libraries in the Big Data era.[80] Xiaodan and Wei referred to the Big Data impact on

the traditional service pattern of academic libraries, highlighting the fact that library services reformation and upgrade can be accomplished.[81] A year later, Gong et al. suggested several ways in which libraries may be creating innovative management and service, facilitating users' utilization of library resources, and integrating systems for optimizing library management processes.[82] Simović presented a Big Data smart library system that has the potential to create new values and data-driven decisions by incorporating multiple sources of differential data.[83] Ye presented a technology framework that included the distributed storage and parallel computing model of mass data, the distributed management model of diverse data sources, and the model of diversified service application for university libraries.[84] Zheng et al. considered university libraries as important supports for teaching and researching in the era of Big Data.[85] Finally, Liu proposed a Big Data application model system of academic library systems based on large scale network analysis method,[86] and Cottrell and Bell presented the case of the Laverne and Dorothy Brown Library at the University of St. Francis (USF), which launched a data tracking initiative via Google Analytics to determine the extent that students use mobile-specific library resources to access library materials and services.[87]

Zeng addressed another important issue relating to the research field of university library management: information ethics. He believed this is an important factor in the era of Big Data that cannot be ignored at present.[88]

All these papers highlight the evolving field of Big Data research in academic libraries, and it seems that academic librarians have a clear role in Big Data analytics by helping library users to enhance the optimization of the library management process. Nevertheless, despite the volume of research, very few studies discussed the implications of analyzing Big Data tools and techniques in academic libraries.

Big Data in Public Libraries

Public libraries can bridge the gap between the general population and the knowledge of what happens with data and how it can be used. This role adds value to the libraries in anticipating needs and then providing community-based services. They can help people better understand data generation and decide in what kind of data world they wish to function.[89] The implementation of Big Data and Big Data technologies in public libraries is explored in three articles. Kim and Cooke conducted a Big Data analysis, which aimed to evaluate library services using the Chernoff faces, a method of visualizing multidimensional data. Big Data in public libraries in London and Seoul were collected, respectively, from the Chartered Institute of Public Finance and Accountancy and the Korean government's website for drawing a Chernoff face.[90] Bertot et al. discussed Local Big Data and the role of libraries in building community data infrastructures. They used a case study approach in a medium-sized U.S. city and discussed the issues that smaller communities face when seeking to create local data infrastructures and the extent to which libraries can develop their capabilities, capacity, and abilities to work with community information and data to facilitate community engagement and high-impact, locally relevant analytics.[91] Finally, Zhan and Widén outlined several roles that public libraries in Finland should undertake in the context of Big Data to better serve citizens, communities and organizations.[92] Therefore, it is obvious that public librarians focus a lot on community-based services. Public libraries serve as cultural centers to their patrons, aiming to provide added value services to them and further benefit the community.

Big Data in Research Libraries

As mentioned by Colby and Levin, librarians should collaborate with researchers in creating new models that support the curation and visualization of Big Data, as well as recommend tools and strategies for data analysis and also advise researchers on data management planning.[93] Three articles attempt to define the role of Big Data within the research libraries context. Furlough believed that librarians should address Big Data in the same way they address any information: for the long term. Nevertheless, the main difference is that, regarding Big Data, librarians are increasingly asked to involve themselves at earlier stages of the information lifecycle.[94] Finally, Federer mentioned that librarians can be valuable collaborators with research teams facing knowledge management challenges: management, analysis, and preservation of research data. He attempted to survey how research has evolved in the age of Big Data and how librarians and other information professionals can respond to researchers' emerging needs.[95] Xie and Fox explored the roles that research libraries should play in the research lifecycle and discuss the motivations behind a research project to investigate effective library Big Data cyberinfrastructure strategies.[96]

These papers present several possibilities and opportunities for the use of Big Data in research libraries, focusing on the analysis, management, and preservation of research data. Literature review indicated that understanding how researchers interact with their data at various points in the lifecycle can help librarians develop targeted services to meet the particular needs of researchers at their institutions and also enhance their services by implementing big data technologies.

Big Data in Government Libraries

Three studies refer to government libraries. Reinhalter and Wittman described the policies regarding the current data surge, including the government-led Open Data Initiative, and discuss the impact on libraries. Data.gov is an interesting example of Big Data, which includes more than 88,400 datasets.[97] In addition, Schwartz mentioned that libraries will be able to identify, select, and acquire large datasets of valuable information content without cost or copyright restrictions, adding value at the same time to this content by preserving the content and providing user-focused services.[98] He referred to Project Open Data, an online, public repository intended to foster collaboration and promote the continual improvement of the Open Data Policy and foster a culture change in government. Finally, Vaidhyanathan and Bulock mentioned that massive governmental data offer librarians an opportunity to discuss the societal implications of "Big Data" and the policies that govern them.[99]

This review reveals that there is little research regarding the management of massive governmental data in the library world. Government librarians should enhance their role in the Big Data era and provide more effective services to the citizens they serve. They admit that there are large potentials as well as large pitfalls while dealing with large governmental data, such as concerns about managing privacy and concerns about the areas of expertise that librarians should enhance to better deal with access and collection management.

Big Data in Medical Libraries

Five papers refer to medical libraries by discussing and analyzing existing problems in health information services. Lu et al. analyzed the existing problems in health information service and discuss the importance and urgency of strengthening the health information services

to the community in the age of Big Data.[100] Two years later, they presented a Big Data age-featured innovative health information services model in university libraries in China through transforming their traditional role, constructing a patient-oriented, network-sharing health information resources bank, and forging full-time and part-time health information literacy service teams.[101] Furthermore, Wang and Cao analyzed the demands and challenges to construct a library personalized subject service platform in Big Data environment that aimed to increase data of information resources cheaply and efficiently.[102] Knapp mentioned that librarians will be able to have an important role as data management experts in their institutions by promoting research tools that play a role in Big Data.[103] Finally, as already mentioned in the Academic Libraries subsection, Gong et al. suggested several ways for libraries to create innovative library management and optimization of their services.[104]

From all the above, it is clear that medical libraries actively focus on Big Data technologies to provide better-coordinated clinical research management services to faculty and clinical researchers. It is essential that they explore Big Data analytics and all emerging technologies to deliver clinical research data management services to the research community they serve.

Big Data in Special Libraries

In the Big Data era, special libraries face some unique challenges as well.[105] Five papers refer to special libraries. Murray believed that special librarians need to watch all technological trends that affect libraries in general, the information industry in particular, as well as their organizations. More specifically, they need to take a leadership role in the field and get involved dynamically in their organization's data strategies.[106] Moreover, Johnson mentioned that, nowadays, librarians must develop additional disciplinary knowledge and data management skills. In particular, he believed that the professional skills of the technical librarian remain relevant and valuable in an era of Big Data.[107] Cervone presented an analytics implementation by following a checklist of eight aspects to be considered in the development of a Big Data and analytics strategy.[108] In addition to the technical issues to be explored, there is a myriad of "softer" issues that must be thoroughly considered. By addressing all the above, an organization can successfully begin Big Data initiatives. Rawlins and Silver explored the relationship between Big Data and the academic business library, as well as existing strategic data resources and their usage trends at Michigan State University.[109] Finally, Teague and Legeros focused on the library services that have to be aligned with the organization's Big Data and analytics goals. This meant that the library's mission statement had to be re-examined and redefined to meet the corporate goals.[110]

It is obvious that special libraries still haven't embraced Big Data tools in their services; nevertheless, it is essential to align provided services with Big Data technology to offer added value to their patrons. In particular, corporate librarians can better manage and maintain all company-owned information resources and associated content, thus helping company staff members with projects by conducting extensive research and archiving data.

Conclusions and Future Research

"Big Data" seems to be a multifaceted and evolving term[111] that has caught global attention and has led to creative brainstorming,[112] changing and transforming rapidly the way we live, work, and think. Our research indicated that Big Data is a huge opportunity for libraries, as it can lead to the creation of new roles for the librarians and the information professionals.

Librarians have always been a vital part of the data and information sector.[113] The main aim, therefore, is to be aware always of the evolving data-based research,[114] to facilitate knowledge creation in communities.

This review indicated that Big Data technologies already play an important role in some types of libraries, such as academic and medical libraries, while some other libraries, such as public and research libraries, take still first steps toward implementing a Big Data strategy. Furthermore, this review showed that librarians, in order to facilitate the research process, have to embrace a role in making big datasets more useful, visible, and accessible by systematizing retrieval methods and introducing new information visualization tools. Therefore, it seems that their role will continue to be one of inspiration and creativity.[115]

Concluding, one thing is for sure: still we do not know the real value of raw data and the extent to which they will change the information landscape, but it is an essential task for all information professionals to explore all the new and powerful possibilities of Big Data, to transform it into actionable knowledge that can be used by decision makers, supporting at the same time information and data analysis literacy,[116] so enhancing libraries' services.

In our future research directions, we will further extend the literature review by analyzing the use of Big Data in each category of libraries separately and examining more topics that influence the implementation of Big Data in libraries, such as legal and policy issues, privacy, and data ownership. The in-depth study of all these elements will certainly lead to more accurate and precise conclusions regarding the ways the library world is influenced by Big Data technologies.

There are other issues that the library and information science community should explore regarding Big Data:

- *Distinguish Big Data:* We, as a library community, do not clearly distinguish between Big Data in the management of the library, Big Data storage and manipulation of research data, and big library operational data that documents patron transactions and usage in academic or corporate libraries.
- *Understand Big Data, clients' data, and privacy:* In addition, there is only passing acknowledgment of the need for subject-specific understanding of Big Data, data of clients, and the privacy that is in conflict with efforts to personalize services and collect data for assessment and reporting.
- *Understand types of data:* Furthermore, how many library staff understand types of scientific data in order to make a meaningful contribution to its storage, use and reuse?
- *Training:* Thus, is there any need to train and provide new skills to library staff concerning this subject area?
- *Services based on Fourth Industrial Revolution:* Finally, how far our library community have been implemented in services regarding Big Data, Artificial Intelligence (AI), the Internet of Things (IoT), and developments that shape the fourth industrial revolution?

It is our duty as information professionals to follow the rapid changes in society, to be well informed and prepared, well trained, work closely with other communities, build synergies, and meet future challenges for the benefit of our society. These issues might set up a future agenda for the library and information science community.

APPENDIX A

TABLE 1
Big Data in Libraries

Reference	Year of Publication	Type of Paper	Keywords	Type of Big Data	Type of Library
Chen, H. et al.	2015	Conference paper	Big data analytics, decision making, data-driven library management	Library transaction logs	Academic libraries
Hoy, M.B.	2014	Journal article	Big data, internet, research data management	Research data	All
Reinhalter, L. and Wittman, R.J.	2014	Journal article	Big data curation, open government, big data policies	Government data	Government libraries
Kim, Y. and Cooke, L.	2017	Journal article	Big data analysis, performance evaluation, data visualization	Library statistical data (staff, collections, budgets, library visits)	Public libraries
Golub, K. and Hansson, J.	2015	Presentation	Big data curation, data visualization	Research data	Academic libraries
Klapwijk, W.	2016	Presentation	Data mining, big data analytics, data visualization	Library statistical data	All
Wang, Ch. et al.	2016	Conference paper	Decision making, library discovery, big data analytics, big data curation	Library data (library records, loan transactions)	All
Goldberg, D. et al.	2014	Journal article	GIS data management, cloud computing	GIS data	GIS data libraries, Academic libraries
Furlough, M.	2012	Presentation	Research data management, policies	Research data	Research libraries
Fister, B.	2015	Journal article	Learning analytics, decision making	Locally generated data, circulation records, electronic book packages, research data	Academic libraries
Bertot, J. et al.	2014	Conference paper	Qualitative analysis, big data analytics, big data curation	Demographic data, voting district-level data	Public libraries
Rawlins, K.G. and Silver, B.	2015	Journal article	Library statistics, big data analytics	Business data, consumer demographic data	Special libraries

TABLE 1
Big Data in Libraries

Reference	Year of Publication	Type of Paper	Keywords	Type of Big Data	Type of Library
Teets, M. and Goldner, M.	2013	Journal article	Big data curation, WorldCat, linked data	Bibliographic data	All
Jantti, M.	2016	Book chapter	Library impact, big data analytics, learning analytics, data warehousing	Library usage data	Academic libraries
Federer, L.	2016	Journal article	Research data management, data sharing, data reuse	Research data	Research libraries
Kumar, N.V.N.P. and Priyadarsini, U.	2016	Journal article	Big data analytics, knowledge creation	Library data, research data	All
Green, R. and Panzer, M.	2014	Journal article	WorldCat, DDC, facet analysis, classification analytics	Bibliographic data	All
Teague, E. and Legeros, J.	2014	Journal article	Big data analytics	Library data, website visits, research requests, reference questions	Special libraries
Cervone, H.F.	2016	Journal article	Big data analytics, analytics project implementation, big data project implementation	Social media postings, computer system logs	Special libraries
He, B. and Zhang, H.	2016	Conference paper	Big data, decision support systems, advanced users	Library data	All
Halperin, M. and Lusk, E.J.	2016	Journal article	Big data analytics, libraries	Library data	Academic libraries
Lu, J. et al.	2016	Journal article	Big data, health information literacy	Health information	Academic libraries Medical libraries
Ping, H.	2015	Journal article	Big data, association rules, personalized recommendation algorithm	Readers' purpose, library collections, library users' interests	All
Allalouf, M.	2015	Conference paper	Big data, data visualization	Library catalog data	Academic libraries
Lu, J.-Y. et al.	2014	Conference paper	Big data, health information services	Health information	Academic libraries Medical libraries

TABLE 1
Big Data in Libraries

Reference	Year of Publication	Type of Paper	Keywords	Type of Big Data	Type of Library
Sun, N.N. and Ma, L.H.	2014	Journal article	Big data, data privacy, users' privacy information security	Library data	All
Sun, C.L. et al.	2014	Journal article	Big data, construction, digital library	Library data, reader behavior data	Academic libraries
Li, Y.	2014	Journal article	Big data analysis, libraries, information resource configuration	Library data	Academic libraries
Wei, C. et al.	2014	Journal article	Big data mining, data analysis, libraries	Library data	All
Yi, C. et al.	2014	Journal article	Big data mining, libraries, clustering analysis	Library data	Academic libraries
Wang, X. and Cao, H.	2014	Conference paper	Big data, library personalized subject service platform, information resources	Library data, subject headings	Medical libraries
Schwartz, M.	2013	Journal article	Big data curation, open government, big data policies	Government data	Government libraries
Campbell, D.G. and Cowan, S.R.	2016	Journal article	Big data analytics, linked data, libraries, privacy	Social media posts	All
Ye, C.	2018	Conference paper	Big data, libraries, computational modeling, data models, parallel processing	Library user data, library service requests	Academic libraries
Liu, Y	2018	Conference paper	Academic libraries, Big data, cloud computing, data analysis, knowledge management	Library data	Academic libraries
Zheng, L. et al.	2018	Conference paper	Big data, university libraries, information security	Network information	Academic libraries
Zhan, M. and Widén, G.	2018	Journal article	Big data, public libraries, professional roles	Collection data, library economic statistics, user information, borrowing histories	Public libraries
Xu, S. et al.	2017	Journal article	Big data, data analysis, data access, knowledge creation	Library data	All

TABLE 1
Big Data in Libraries

Reference	Year of Publication	Type of Paper	Keywords	Type of Big Data	Type of Library
Linlin, Z.	2017	Journal article	Big data, virtual libraries, experiential teaching	Library data	Academic libraries
Golub, K. and Hansson, J.	2017	Journal article	Big data, scholarly communication, bibliometrics, data sharing, data curation	Library data, open data, research data, GIS data.	All
Ahmed, W. and Ameen, K.	2017	Journal article	Big data, library and information management	Library data, geospatial data	All
Xiaodan, D. and Wei, W.	2017	Conference paper	Big data, libraries, service pattern	Library data	Academic libraries
Long, S.	2018	Conference paper	Big data, digital libraries, information services, resource allocation	Library data	All
Olendorf, R., Wang, Y.	2017	Book chapter	Big data, libraries, needs analysis, visualization	Library data, research data	All
Bhat, W.A.	2018	Journal article	Digital libraries, digital preservation, big data, data reduction, magnetic storage	Digital library data	All
Gong, R. et al.	2018	Journal article	Statistical education, big data relationship quality, innovative service of library	Membership database of the library	Academic libraries Medical libraries
Chang, C.C.	2018	Journal article	Big data, library services, genealogical information system, migration analysis, spatial-temporal visualization	Genealogical information, GIS data	GIS data libraries
Hao, W.	2017	Journal article	Big data, digital storage, information services, systems analysis	Library data	All
Qin, S.	2018	Journal article	Big data, data mining information management	Library data	All
Simović, A.	2018	Journal article	Libraries, data analysis, Big data, data storage, recommender system	Data from different sources	Academic libraries
Li, J. et al.	2017	Journal article	Libraries, big data, feasibility analysis, personalized information push service	Library data	All

TABLE 1
Big Data in Libraries

Reference	Year of Publication	Type of Paper	Keywords	Type of Big Data	Type of Library
Xie, Z., Fox, E.A.	2017	Journal article	Data management, big data, cyberinstrastructure, data sharing, data reuse	Research data	Research libraries
Luo, Y.C. et al.	2013	Journal article	Big data, smart libraries, RFID technology, cloud technology	Library data	Academic libraries
Cervone, H.F.	2017	Journal article	Social media analytics, social media presence, big data social media evaluation, social network analysis	Social media posts	All
Murray, T.E.	2016	Journal article	Big data, special libraries, open access, open data, information-seeking behavior	Network information	Special libraries
Cottrell, T. and Bell, B.	2016	Journal article	Big data, mobile technology, data analytics, libraries	Mobile library data	Academic libraries
Deng, Z.	2017	Conference paper	Service innovation, academic libraries, big data	Network information	Academic libraries
Gerrard et al.	2018	Journal article	Big data, digital preservation, image processing, business analytics, digitization	Research data	Academic libraries
Harper and Oltmann	2017	Bulletin article	Big data, libraries, privacy, user behavior, social media	Social media posts	All
Knapp, M.	2013	Journal article	Big data, medical informatics, medical libraries, health science	Information from searching habits, purchasing trends and demographic information	Medical libraries
Kong, N.N.	2015	Journal article	GIS data, big data management, GIS libraries	GIS data	GIS data libraries Academic libraries
Leetaru, K.	2016	Conference paper	Big data, libraries, social media	Social media posts	All

TABLE 1
Big Data in Libraries

Reference	Year of Publication	Type of Paper	Keywords	Type of Big Data	Type of Library
Zhan, M. and Widén, G.	2017	Journal article	Big data, definition analysis, librarian skills, librarianship	Library data (data collection, borrowing and lending transactions, user profiles), social media data, data from mobile devices	All
Poole and Garwood	2018	Journal article	Digital humanities, digital curation, data curation, information and library science	Data from digital humanities projects	All
Tuppen, S. et al.	2016	Journal article	Big data, music data, libraries	Library catalogue records	Academic libraries
Vaidhyanathan, S. and Bulock, Ch.	2014	Journal article	Big data, privacy, social engagement	Massive governmental data	Government libraries
Wang, D.X.	2016	Conference paper	Big data, information service, academic libraries	Library data	Academic libraries
Zeng, Y.	2015	Journal article	Big data, libraries, information ethics, standard construction	Network information	Academic libraries
Johnson, V.	2017	Journal article	Database development, GIS data, special libraries, embedded librarians, cross-disciplinary collaboration	GIS data	Special libraries
Weessies, Kathleen W., and Daniel S. Dotson	2013	Journal article	Data visualization, big data	GIS data	Academic libraries

APPENDIX B

TABLE 2
Identified Papers' Type of Publication

Kind of Paper	No. Papers	Papers
Journal article	49	Hoy (2014), Reinhalter and Wittman (2014), Kim and Cooke (2017), Goldberg et al. (2014), Fister (2015), Rawlins and Silver (2015), Teets and Goldner (2013), Federer (2016), Kumar and Priyadarsini (2016), Green and Panzer (2014), Teague and Legeros (2014), Cervone (2016), Halperin and Lusk (2016), Lu et al. (2016), Ping (2015), Sun and Ma (2014), Sun et al. (2014), Li (2014), Wei et al. (2014), Yi et al. (2014), Schwartz (2013), Campbell and Cowan (2016), Zhan and Widén (2018), Xu et al. (2017), Linlin (2017), Golub and Hansson (2017), Ahmed and Ameen (2017), Hao (2017), Li et al. (2017), Xie and Fox (2017), Bhat (2018), Gong et al. (2018), Chang (2018), Qin (2018), Simović (2018), Luo et al. (2013), Cervone (2017), Murray (2016), Cottrell and Bell (2016), Gerrard et al. (2018), Knapp (2013), Kong (2015), Zhan and Widén (2017), Poole and Garwood (2018), Tuppen et al. (2016), Vaidhyanathan and Bulock (2014), Zeng (2015), Johnson (2017), Weessies and Dotson (2013)
Presentation	3	Golub and Hansson (2015), Kalpwijk (2016), Furlough (2012)
Book chapter	2	Jantti (2016), Olendorf and Wang (2017)
Conference paper	15	Chen et al. (2015), Wang et al. (2016), Bertot et al. (2014), He and Zhang (2016), Allalouf (2015), Lu et al. (2014), Wang and Cao (2014), Ye (2018), Liu (2018), Zheng et al. (2018), Xiaodan and Wei (2017), Long (2018), Deng (2017), Leetaru (2016), Wang (2016)
Bulletin article	1	Harper and Oltmann (2017)

TABLE 3
Number of Identified Relevant Articles per Year

Year	No. Papers	Papers
2012	1	Furlough (2012)
2013	5	Teets and Goldner (2013), Schwartz (2013), Luo et al. (2013), Knapp (2013), Weessies and Dotson (2013)
2014	14	Hoy (2014), Reinhalter and Wittman (2014), Goldberg et al. (2014), Bertot et al. (2014), Green and Panzer (2014), Teague and Legeros (2014), Lu et al. (2014), Sun and Ma (2014), Sun et al. (2014), Li (2014), Wei et al. (2014), Yi et al. (2014), Wang and Cao (2014), Vaidhyanathan and Bulock (2014)
2015	8	Chen et al. (2015), Golub and Hansson (2015), Fister (2015), Rawlins and Silver (2015), Ping (2015), Allalouf (2015), Kong (2015), Zeng (2015)
2016	15	Kalpwijk (2016), Wang et al. (2016), Jantti (2016), Federer (2016), Kumar and Priyadarsini (2016), Cervone (2016), He and Zhang (2016), Halperin and Lusk (2016), Lu et al. (2016), Campbell and Cowan (2016), Murray (2016), Cottrell and Bell (2016), Tuppen et al. (2016), Wang (2016), Leetaru (2016)
2017	15	Kim and Cooke (2017), Olendorf and Wang (2017), Xu et al. (2017), Linlin (2017), Golub and Hansson (2017), Ahmed and Ameen (2017), Xiaodan and Wei (2017), Hao (2017), Li et al. (2017), Xie and Fox (2017), Cervone (2017), Zhan and Widén (2017), Johnson (2017), Deng (2017), Harper and Oltmann (2017)
2018	12	Ye (2018), Liu (2018), Zheng et al. (2018), Zhan and Widén (2018), Long (2018), Bhat (2018), Gong et al. (2018), Chang (2018), Qin (2018), Simović (2018), Gerrard et al. (2018), Poole and Garwood (2018)

TABLE 4
Library Type of Identified Relevant Papers

Type of Library	No. Papers	Papers
Academic libraries	28	Chen et al. (2015), Golub and Hansson (2015), Goldberg et al. (2014), Fister (2015), Jantti (2016), Halperin and Lusk (2016), Lu et al. (2016), Sun et al. (2014), Li (2014), Yi et al. (2014), Lu et al. (2014), Allalouf (2015), Ye (2018), Liu (2018), Zheng et al. (2018), Linlin (2017), Xiaodan and Wei (2017), Gong et al. (2018), Simović (2018), Luo et al. (2013), Weessies and Dotson (2013), Gerrard et al. (2018), Deng (2017), Kong (2015), Cottrell and Bell (2016), Tuppen et al. (2016), Wang (2016), Zeng (2015)
Public libraries	3	Kim and Cooke (2017), Bertot et al. (2014), Zhan and Widén (2018)
Research libraries	3	Furlough (2012), Federer (2016), Xie and Fox (2017)
GIS data libraries	3	Goldberg et al. (2014), Chang (2018), Kong (2015)
Government libraries	3	Reinhalter and Wittman (2014), Schwartz (2013), Vaidhyanathan and Bulock (2014)
Medical libraries	5	Lu et al. (2016), Wang and Cao (2014), Lu et al. (2014), Gong et al. (2018), Knapp (2013)
Special libraries	5	Rawlins and Silver (2015), Teague and Legeros (2014), Cervone (2016), Murray (2016), Johnson (2017)
All	25	Hoy (2014), Kalpwijk (2016), Wang et al. (2016), Teets and Goldner (2013), Kumar and Priyadarsini (2016), Green and Panzer (2014), Campbell and Cowan (2016), Wei et al. (2014), Sun and Ma (2014), He and Zhang (2016), Ping (2015), Xu et al. (2017), Golub and Hansson (2017), Ahmed and Ameen (2017), Long (2018), Olendorf and Wang (2017), Bhat (2018), Hao (2017), Qin (2018), Li et al. (2017), Cervone (2017), Harper and Oltmann (2017), Leetaru (2016), Zhan and Widén (2017), Poole and Garwood (2018)

Notes

1. Lauren Reinhalter and Rachel Jane Wittman, "The Library: Big Data's Boomtown," *Serials Librarian: From the Printed Page to the Digital Age* 67, no. 4 (2014): 363–72.

2. Reinhalter and Wittman, "The Library: Big Data's Boomtown," 363–72.

3. Matthew B. Hoy, "Big Data: An Introduction for Librarians," *Medical Reference Services Quarterly* 33, no. 3 (July 2014): 320–26.

4. Mihyun Chung and Jaehyoun Kim, "The Internet Information and Technology Research Directions Based on the Fourth Industrial Revolution," *KSII Transactions on Internet and Information Systems* 10 (March 2016): 1311–20.

5. J. Bloem et al., "The Fourth Industrial Revolution. Things to Tighten the Link between IT and OT". (2014); Keliang Zhou, Taigang Liu, and Lifeng Zhou, "Industry 4.0: Towards Future Industrial Opportunities and Challenges," *12th International Conference on Fuzzy Systems and Knowledge Discovery (FSKD)* (2015); Michael A. Peters, "Technological Unemployment: Educating for the Fourth Industrial Revolution," *Journal of Self-Governance and Management Economics* 5, no. 1 (2017): 32–41; Okyay Kaynak and Shen Yin, "Big Data for Modern Industry: Challenges and Trends [Point of View]," *Proceedings of the IEEE* 103, no. 2 (2015):143–46; H. Kagermann, W. Wahlster, and W. Wahlster, "Securing the Future of German Manufacturing Industry: Recommendations for Implementing the Strategic Initiative INDUSTRIE 4.0," *Final Report of the Industrie 4.0 Working Group* (2013); D. Ellen Frederick, "Libraries, Data and the Fourth Industrial Revolution (Data Deluge Column)," *Library Hi Tech News* 33, no. 5

(2016): 9–12.

6. Sirje Virkus and Emmanouel Garoufallou, "Data science and its relationship to library and information science: a content analysis", *Data Technologies and Applications*, 54 no. 5 (2020): 643-663, https://doi.org/10.1108/DTA-07-2020-0167; Sirje Virkus and Emmanouel Garoufallou, "Data Science from a Library and Information Science Perspective," *Data Technologies and Applications* 53, no. 4 (2019): 422–41, https://doi.org/10.1108/DTA-05-2019-0076; Sirje Virkus and Emmanouel Garoufallou, "Data Science from a Perspective of Computer Science," in *Metadata and Semantic Research*, eds. E. Garoufallou, F. Fallucchi, and E. William De Luca, MTSR 2019, Communications in Computer and Information Science, vol. 1057 (Cham, Switzerland: Springer, 2019), https://doi.org/10.1007/978-3-030-36599-8_19.

7. Zhaoyan Deng, "Research on Service Innovation of Library in Big Data Age," *Proceedings of the 4th International Conference on Education Management and Computing Technology* (ICEMCT 2017, Hangzhou, China, April 15–16, 2017), 1014–17.

8. Koraljka Golub and Joacim Hansson, "Big Data in Library and Information Science," *Big Data: Från Hype till Handling*, Linnaeus University, December 2015, https://www.diva-portal.org/smash/get/diva2:885856/FULLTEXT01.pdf [accessed 29 August 2018].

9. Barbara Blummer and Jeffrey M. Kenton, "Big Data and Libraries: Identifying Themes in the Literature," *Internet Reference Services Quarterly* 1 (January 2019): 1–26.

10. Deng, "Research on Service Innovation of Library in Big Data Age."

11. Dong-Xia Wang, "Information Service in the Big Data Era and Development Strategies for University Libraries," *Proceedings of 2nd Annual International Conference on Social Science and Contemporary Humanity Development* (SSCHD 2016, Wuhan, Hubei, China, July 15–17, 2016): 664–68.

12. Shoban Babu Sriramoju, *Introduction to Big Data: Infrastructure and Networking Considerations* (Cadillac, MI: Horizon Books, 2017).

13. Doug Laney, "3-D Data Management: Controlling Data Volume, Velocity and Variety," *META Research Note* (February 2001): 6

14. Edd Dumbill, "Making Sense of Big Data," *Big Data* 1, no. 1 (February 2013): 1–2.

15. Reinhalter and Wittman, "The Library: Big Data's Boomtown."

16. Noam Slonim et al., "Knowledge-Analytics Synergy in Clinical Decision Support," *Studies in Health Technology and Informatics* 180 (2012): 703–07.

17. Michael Schroeck et al., "Analytics: The Real-world Use of Big Data: How Innovative Enterprises Extract Value from Uncertain Data," *IBM Global Business Services, Business Analytics and Optimization, Executive Report*, 2012.

18. Andrea De Mauro, Marco Greco, and Michele Grimaldi, "A Formal Definition of Big Data Based on Its Essential Features," *Library Review* 65, no. 3 (2016): 122–35.

19. Wouter Klapwijk, "The Library (Big) Data Scientist," *Big Data: New Roles and Opportunities for New Librarians* (IFLA New Professionals Special Interest Group Webinars International Federation of Library Associations and Institutions, 2016).

20. Stephen Mutula, "Big Data Industry: Implication for the Library and Information Sciences," *African Journal of Library Archives and Information Science* 26, no. 2 (October 2016): 93–96.

21. Reinhalter and Wittman, "The Library: Big Data's Boomtown."

22. Panorea Gaitanou, Emmanouel Garoufallou, and Panos Balatsoukas, "The Effectiveness of Big Data in Health Care: A Systematic Review," in *Metadata and Semantics Research*, eds. Sissi Closs et al. (Cham, Switzerland: Springer, 2014), Proceedings of 8th Research Conference, MTSR 2014, Karlsruhe, Germany, November 27–29, 2014, *Communications in Computer and Information Science (CCIS)* 478 (2014): 141–53, https://doi.org/10.1007/978-3-319-13674-5_14.

23. Ron Winslow, "'Big Data' for Cancer Care: Vast Storehouse of Patient Records Will Let Doctors Search for Effective Treatment," *Wall Street Journal*, updated March 26, 2013, https://www.wsj.com/articles/SB10001424127887323466204578384732911187000 [accessed 30 August 2018]; Cota Team, "The Promise of Big Data for Cancer Patients and Practices," https://www.cotahealthcare.com/post/how-cancer-practices-and-patients-benefit-from-big-data [accessed 21 August 2018].

24. "Metadata and Semantics Research", eds. E. Garoufallou and M. Ovalle-Perandones. (Cham, Switzerland: Springer, 2021). 14th International Conference, MTSR 2020, Madrid, Spain, December 2-4, 2020), Revised Selected Papers, *Communications in Computer and Information Science* (CCIS) 1355 (2021): ISBN 978-3-030-71902-9 (in press); "Metadata and Semantics Research", eds. E. Garoufallou et al. (Cham, Switzerland: Springer, 2019). 13th International Conference, MTSR 2019,Rome, Italy, October 28–31, 2019), Revised Selected Papers, *Communications in Computer and Information Science* (CCIS) 1057 (2019): 1-460, https://doi.org/10.1007/978-3-030-36599-8; "Metadata and Semantic Research," eds. E. Garoufallou et al. (Cham, Switzerland: Springer, 2018), 12th International

Conference, MTSR 2018, Limassol, Cyprus, October 23–26, 2018), Revised Selected Papers, *Communications in Computer and Information Science (CCIS)* 846 (2019): 1–378, https://doi.org/10.1007/978-3-030-14401-2; "Metadata and Semantics Research," eds. E. Garoufallou et al. (Cham, Switzerland: Springer, 2017), Proceedings of 11th International Conference, MTSR 2017, Tallinn, Estonia, November 28–December 1, 2017, *Communications in Computer and Information Science (CCIS)* 755 (2017): 1–334, https://doi.org/10.1007/978-3-319-70863-8; "Metadata and Semantics Research," eds. E. Garoufallou et al. (Cham, Switzerland: Springer, 2016), Proceedings of 10th International Conference, MTSR 2016, Göttingen, Germany, November 22-25, 2016, *Communications in Computer and Information Science (CCIS)* 672 (2016): 1–378, https://doi.org/10.1007/978-3-319-49157-8.

25. Sheila Corrall, "Big Data 2.0: Critical Roles for Libraries and Librarians," *Proceedings of the 36th Annual Charleston Conference: Issues in Book and Serial Acquisition* (Charleston, SC, October 31–November 6, 2016).

26. Hoy, "Big Data: An Introduction for Librarians."

27. Hoy, "Big Data: An Introduction for Librarians."

28. Joe McKendrick, "3 Reasons Why 'Big Data' Can Often Be Meaningless or Misleading," ZDNet (2013), www.zdnet.com/article/3-reasons-why-big-data-can-often-be-meaningless-or-misleading [accessed 12 August 2018].

29. Hoy, "Big Data: An Introduction for Librarians."

30. Golub and Hansson, "Big Data in Library and Information Science."

31. Victoria Louise Lemieux, Brianna Gormly, and Lyse Rowledge, "Meeting Big Data Challenges with Visual Analytics: The Role of Records Management," *Records Management Journal* 24, no. 2 (2014): 122–41.

32. Pippa Hemingway and Nic Brereton, "What Is a Systematic Review?" *Evidence-Based Medicine*, 2nd ed. (2009).

33. Shuqing Li et al., "Problems and Changes in Digital Libraries in the Age of Big Data from the Perspective of User Services," *Journal of Academic Librarianship* 45, no.1 (2019): 22–30; Jialing Zhao, Wenwei Cai, and Xiangyuan Zhu, "Research on Smart Library Big Service Application in Big Data Environment," in *Advances in Computational Science and Computing*, eds. Ning Xiong et al. (Cham, Switzerland: Springer, 2019), 238–45; Grzegorz Osinski, Veslava Osinska, and Piotr Malak, "PCA Algorithms in the Visualization of Big Data from Polish Digital Libraries," in *Multimedia and Network Information Systems* (Advances in Intelligent Systems and Computing), eds. K. Choroś et al. (Cham, Switzerland: Springer, 2019), 833; Blummer and Kenton, "Big Data and Libraries: Identifying Themes in the Literature."

34. Ming Zhan and Gunilla Widén, "Understanding Big Data in Librarianship," *Journal of Librarianship and Information Science* 1, no. 16 (2017): 561–76.

35. Hoy, "Big Data: An Introduction for Librarians."

36. N.V.N. Praveen Kumar and Uma Priyadarsini, "Revealing Library Statistics with Big Data Expertise: A Review," *International Journal of Pharmacy and Technology* 8, no. 4 (December 2016): 20783–89.

37. Shaochun Xu et al., "The Library Big Data Research: Status and Directions," *International Journal of Software Innovation* 5, no. 3 (2017): 77–88.

38. Lindsey M. Harper and Shannon M. Oltmann, "Big Data's Impact on Privacy for Librarians and Information Professionals," *Bulletin of the Association for Information Science and Technology* 43, no. 4 (2017): 19–23.

39. Kalev Leetaru, "Reimagining Our World at Planetary Scale: The Big Data Future of Our Libraries," *Proceedings of the 36th Annual Charleston Conference: Issues in Book and Serial Acquisition* (Charleston, SC, October 31–November 6, 2016), 58–62.

40. Klapwijk, "The Library (Big) Data Scientist."

41. Chunning Wang et al., "Exposing Library Data with Big Data Technology: A Review," *IEEE/ACIS 15th International Conference on Computer and Information Science* (ICIS, Okayama, Japan, June 26–29, 2016), 1–6.

42. Jun Li et al., "Big Data Application Framework and in Library," *Information Discovery and Delivery* 45, no. 4 (2017): 161–68.

43. Waqar Ahmed and Kanwal Ameen, "Defining Big Data and in the Field of Information and Library Management," *Library Hi Tech News* 34, no. 9 (2017): 21–24.

44. H. Frank Cervone, "Organizational Considerations Initiating a Big Data and Analytics Implementation," *Digital Library Perspectives* 32, no. 3 (2016): 137–41.

45. Shuai Qin, "The Application of Big Data Thinking in Library Information Management System," *Journal of Advanced Oxidation Technologies* 21, no. 2 (2018).

46. Wang Hao, "Personalized Information Service System of a Library under the Big Data Environment," *Agro Food Industry Hi-Tech* 28, no. 1 (January 2017): 1701–04.

47. Wasim Ahmad Bhat, "Long-term Preservation of Big Data: Prospects of Current Storage Technologies in Digital Libraries," *Library Hi Tech* 36, no. 3 (2018): 539–55.

48. Koraljka Golub and Joacim Hansson, "(Big) Data in Library and Information Science: A Brief Overview

of Some Important Problem Areas," *Journal of Universal Computer Science* 23, no. 11 (2017): 1098–1108.

49. Suyan Long, "Information Service Research and Development of Digital Library in the Era of Big Data," *Proceedings of 2017 13th International Conference on Semantics, Knowledge and Grids* (SKG 2017, January 2018), 150–53.

50. Robert Olendorf and Yan Wang, "Big Data in Libraries," in *Big Data and Visual Analytics*, eds. S. Suh and T. Anthony (Cham, Switzerland: Springer International Publishing, 2018), 191–202.

51. Ning Ning Sun and Li Hua Ma, "Research on Information Security and Privacy of Libraries in Big Data Era," *Advanced Materials Research* 1049/1050 (October 2014): 1934–37.

52. Chen Wei et al., "Analysising on the Library Services in the Age of Big Data," *Advanced Materials Research* 1044/1045 (October 2014): 1066–70.

53. Bo He and Hongyuan Zhang, "Library Personalized Information Recommendation of Big Data," *Proceedings of 2016 IEEE International Conference of Online Analysis and Computing Science* (Chongqing, China, 2016), 289–92.

54. He Ping, "The Research on Personalized Recommendation Algorithm of Library Based on Big Data and Association Rules," *Open Cybernetics and Systemics Journal* 9 (2015): 2554–58.

55. Grant D. Campbell and Scott Cowan, "The Paradox of Privacy: Revisiting a Core Library Value in an Age of Big Data and Linked Data," *Library Trends* 64, no. 3 (2016): 492–511.

56. Alex H. Poole and Deborah A. Garwood, "Natural Allies" Librarians, Archivists, and Big Data in international digital humanities project work," *Journal of Documentation* 74, no. 4 (2018): 804–26.

57. Michael Teets and Matthew Goldner, "Libraries' Role in Curating and Exposing Big Data," *Future Internet* 5 (2013): 429–38.

58. Rebeca Green and Michael Panzer, "The Interplay of Big Data, WorldCat, and Dewey," *Advances in Classification Research Online* 24, no. 1 (2014): 51–58.

59. Hsin-liang Chen et al., "Library Assessment and Data Analytics in the Big Data Era: Practice and Policies," *Proceedings of the 78th ASIST Annual Meeting: Information Science with Impact: Research in and for the Community* (St. Louis, Missouri, November 6–10, 2015), 1–4.

60. Chen et al., "Library Assessment and Data Analytics in the Big Data Era."

61. Deng, "Research on Service Innovation of Library in Big Data Age."

62. Wang, "Information Service in the Big Data Era and Development Strategies for University Libraries."

63. Yong Cheng Luo, Jing Jing Cao, and Jing Fang Qian, "Exploration and Construction of Smart Library Based on RFID Technology," *Advanced Materials Research* 765/767(2013): 1743–46.

64. David Maynard Gerrard, James Edward Mooney, and Dave Thompson, "Digital Preservation at Big Datascales: Proposing a Step-change in Preservation System Architectures," *Library Hi Tech* 36, no. 3 (2018): 524–38.

65. Jianyu Lu et al., "Establishing a University Library-based Health Information Literacy Service Model in the Age of Big Data," *Journal of Medical Imaging and Health Informatics* 6, no. 1 (February 2016): 260–63; Jianyu Lu et al., "Thinking on Strengthening University Library Health Information Service Model Innovation in the Age of Big Data," in *Medicine Sciences and Bioengineering: Proceedings of the 2014 International Conference on Medicine Sciences and Bioengineering (ICMSB2014)*, ed. Mings Wang (Kunming, Yunnan, China, August 16–17, 2014), 603–06.

66. C. Yi, Y. Xia, and Z.Y. Zhang, "Study on the Personal Push Service of University Library Based on Big Data Mining," *Advanced Materials Research* 1, no. 1 (2014): 1261–65.

67. Chang Li Sun, Dan Shi, and Dong Mei Guo, "Analysis on the Library Digitization Construction under the Big Data Era of University," *Applied Mechanics and Materials* 687/691 (November 2014): 2656–59.

68. Miriam Allalouf, "Big Data in the Library: Extending Modern Library Catalogues with Data Visualization, Linking and Mining," in *SKY 2015: Proceedings of the 6th International Workshop on Software Knowledge*, in conjunction with IC3K 2015, 74–75.

69. Chen et al., "Library Assessment and Data Analytics in the Big Data Era."

70. Golub and Hansson, "Big Data in Library and Information Science."

71. Barbara Fister, "Big Data or Big Brother? Data, Ethics, and Academic Library," *Library Issues: Briefings for Faculty and Administrators* 35, no. 4 (March 2015).

72. Margie Jantti, "Libraries and Big Data: A New View on Impact and Affect," in *Quality and the Academic Library: Reviewing, Assessing and Enhancing Service Provision*, ed. J. Atkinson (London, UK: Elsevier, 2016), 267–73.

73. Yang Li, "Analysis of the Construction of the Library Information Resources from the Perspective of Big Data," *Applied Mechanics and Materials* 631/632 (September 2014): 1067–70.

74. Daniel Goldberg et al., "Maps & GIS Data Libraries in the Era of Big Data and Cloud Computing," *Journal of Map & Geography Libraries: Advances in Geospatial Information, Collections & Archives* 10, no. 1 (April 2014): 100–22.

75. Nicole Ningning Kong, "Exploring Best Management Practices for Geospatial Data in Academic Libraries," *Journal of Map and Geography Libraries* 11, no. 2 (2015): 207–25.

76. Kathleen W. Weessies and Daniel S. Dotson, "Mapping for the Masses: GIS Lite and Online Mapping

Tools in Academic Libraries," *Information Technology and Libraries* 32, no. 1 (2013): 23–35.

77. Chen Chi Chang, "Hakka Genealogical Migration Analysis Enhancement Using Big Data on Library Services," *Library Hi Tech* 36, no. 3 (2018): 426–42.

78. Michael Halperin and Edward J. Lusk, "Navigating the Benford Labyrinth: A Big-data Analytic Protocol Illustrated Using the Academic Library Context," *Knowledge Management and E-Learning* 8, no. 1 (March 2016): 138–57.

79. Sandra Tuppen, Stephen Rose, and Loukia Drosopoulou, "Library Catalogue Records as a Research Resource: Introducing 'A Big Data History of Music'," *Fontes Artis Musicae* 63, no. 2 (2016): 67–88.

80. Zhang Linlin, "Analysis of the Information Processing Technology of University Libraries in the Big Data Era," *Agro Food Industry Hi-Tech* 28, no. 1 (January 2017): 2036–40.

81. Du Xiaodan and Wang Wei, "Discussion on University Library Service Pattern in Big Data Era," *2017 IEEE 2nd International Conference on Big Data Analysis (ICBDA)* (Beijing, China, 10–12 March 2017).

82. Ruiyi Gong, Kaijun Yu, and Hongmei Tang, "Based on Statistical Education to Study Innovative Service and Relationship Quality of University Library under Big Data," *Eurasia Journal of Mathematics, Science and Technology Education* 14, no. 6 (April 2018): 2419–25.

83. Aleksandar Simović, "A Big Data Smart Library Recommender System for an educational Institution," *Library Hi Tech* 36, no. 3 (2018): 498–523.

84. Chunlei Ye, "Research on the Key Technology of Big Data Service in University Library," *13th International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery* (Guilin, China, 2018): 2573–78.

85. Likun Zheng et al., "A Research on Information Security of University Libraries in the Era of Big Data," in *Recent Developments in Mechatronics and Intelligent Robotics*, eds. F. Qia., S. Patnaik, and J. Wang, ICMIR 2017, Advances in Intelligent Systems and Computing, 690 (Cham, Switzerland: Springer, 2018), 112–16.

86. Yueyun Liu, "Research on the Application of Big Data in Academic Libraries," *Proceedings of 3rd International Conference on Intelligent Transportation, Big Data and Smart City (ICITBS)* (January 2018), 364–67.

87. Terry Cottrell and Brigitte Bell, "When to Say When: Using Big Data to Support Mobile Communication," *College and Undergraduate Libraries* 23, no. 3 (2016): 315–-22.

88. Ying Zeng, "On the Analysis of Library Information Ethics and the standard construction in the Era of Big Data," *Studies in Literature and Language* 11, no. 4 (2015): 25–28.

89. Christina Wilson, "Why Big Data Matters to Libraries," INALJ 2014, http://inalj.com/?p=65209 [accessed 29 August 2018].

90. Young-Seo Kim and Louise Cooke, "Big Data Analysis of Public Library Operations and Services by Using the Chernoff Face Method," *Journal of Documentation* 73, no. 3 (2017): 466–80.

91. John Carlo Bertot, Brian S. Butler, and Diane M. Travis, "Local Big Data: The Role of Libraries in Building Community Data Infrastructures," *Proceedings of the 15th Annual International Conference on Digital Government Research* (Aguascalientes, Mexico, June 18–21, 2014), 17–23.

92. Ming Zhan and Gunilla Widén, "Public Libraries: Roles in Big Data," *Electronic Library* 36, no. 1 (2018): 133–45.

93. Jeanette Colby and Yisrael Levin, "Big Library, Big Data," University of Rochester, www.rochester.edu/newscenter/librarians-integral-to-data-science-initiatives-233042 [accessed 29 August 2018].

94. Mike Furlough, "Research Libraries and "Big Data," *Penn State University Libraries CENDI/NFAIS Workshop* (Washington, December 11, 2012).

95. Lisa Federer, "Research Data Management in the Age of Big Data: Roles and Opportunities for Librarians," *Information Services and Use* 36, no. 1/2 (September 2016): 35–43.

96. Zhiwu Xie and Edward A. Fox, "Advancing Library Cyberinfrastructure for Big Data Sharing and Reuse," *Information Services and Use* 37, no. 3 (November 2017): 319–23.

97. Reinhalter and Wittman, "The Library: Big Data's Boomtown."

98. Meredith Schwartz, "What Governmental Big Data May Mean for Libraries," *Library Journal* 138, no. 10 (May 2013).

99. Silva Vaidhyanathan and Chris Bulock, "Knowledge and Dignity in the Era of 'Big Data'," *Serials Librarian* 66, no. 1/4 (2014): 49–64.

100. Lu et al, "Thinking on Strengthening University Library Health Information Service Model Innovation in the Age of Big Data."

101. Lu et al, "Establishing a University Library-based Health Information Literacy Service Model in the Age of Big Data."

102. Xia Wang and Hongxin Cao, "Study on Building Library Personalized Subject Service Platform in Big Data Environment: Lib 2.0 Solutions Based on Hadoop Framework," *Proceedings of 2014 IEEE 3rd International Conference on Cloud Computing and Intelligence Systems* (Shenzhen, China, 27–29 November 2014), 146–49.

103. Maureen Knapp, "Big Data," *Journal of Electronic Resources in Medical Libraries* 10, no. 4 (2013): 215–22.
104. Gong et al., "Based on Statistical Education to Study Innovative Service and Relationship Quality of University Library under Big Data."
105. David Bender, "What's Special about Special Libraries? SLA and the Continuous Education Challenge," *INSPEL* 32, no. 4 (1998): 197–204.
106. Tara E. Murray, "The Forecast for Special Libraries," *Journal of Library Administration* 56, no. 2 (2016): 188–98.
107. Vanessa Johnson, "Leveraging Technical Library Expertise for Big Data Management," *Journal of the Australian Library and Information Association* 66, no. 3 (2017): 271–86.
108. H. Frank Cervone, "Evaluating Social Media Presence: A Practical Application of Big Data and Analytics in Information Organizations," *Digital Library Perspectives* 33, no. 1 (2017): 2–7.
109. Kara Gust Rawlins and Breezy Silver, "Too 'Big' to Process: Big Data in the Business Library: Key Resources and Implications for the Future," *Issues in Information Systems* 16, no. 3 (2015): 64–69.
110. Elaine Teague and Julie Legeros, "Big Data's Role in Information-Centric Organizations," *Information Outlook* 18, no. 3 (2014): 18–20.
111. Chen et al., "Library Assessment and Data Analytics in the Big Data Era."
112. Terence K. Huwe, "Big Data and the Library: A Natural Fit," *Computers in Libraries* 34, no. 2 (March 2014): 17–18.
113. Reinhalter and Wittman, "The Library: Big Data's Boomtown."
114. Reinhalter and Wittman, "The Library: Big Data's Boomtown."
115. Nathan Rinne, "Big Data, Big Libraries, Big Problems? The 2014 LibTech Anti-talk?" in *Library Technology Conference* (Macalester College, St. Paul, Minnesota, March 19 and 20, 2014).
116. Golub and Hansson, "Big Data in Library and Information Science."