



2018

Peer Reviewed Referred
and UGC Listed Journal
Journal No.: 47100



**AN INTERNATIONAL MULTIDISCIPLINARY
HALF YEARLY RESEARCH JOURNAL**

GENIUS

Volume - VI, Issue - II, FEBRUARY - JULY - 2018
ISSN - 2279 - 0489

Impact Factor - 4.954 (www.sjifactor.com)

PART - III

20

AJANTA PRAKASHAN



CONTENTS OF PART - III



Sr. No.	Name & Author Name	Page No.
12	Multimedia Data Mining Ms. Krutika H. Churi	78-83
13	Green Computing "Eco-Friendly Technology" Ms. Tejal R. Patil	84-88
14	Green Computing-E-Waste Minimization Vaishali Sindekar Yugandhara More	89-93
15	Intellectual Properties Mr. Raut Mrudul Ashok	94-102
16	IOT based Smart Switches Ms. Priyadharshini Thevar Mr. Shiba Prasad Kar	103-107
17	Cyber Security - Problems and Solutions Ms. Manali B. Churi Ms. Niyati S. Patil	108-114
18	Cyber Pornography Mr. Raut Mrudul Ashok	115-119
19	Parallel Database System and Query Evaluation Ms. Kajal Singh Ms. Manali Patil	120-125
20	Strengthening of Cognitive Computing Technology for Students to Improve the Effectiveness in Education Mrs. Priyanka Bangar Mrs. Dipika P. Vishe	126-130
21	MA-NET is a New Pattern of Wireless Communication for the Performance of Mobile Hosts in the Network Dr. Bhanu Pratap	131-139
22	Big Data in Libraries: Challenges and Issues Prof. Sheela K. Godbole Dr. Ramdas Lihitkar	140-146

Big Data in Libraries: Challenges and Issues

Prof. Sheela K. Godbole

Librarian, Sonopant Dandekar Arts, V. S. Apte Commerce & M. H. B. College,
Tal & Dist-Palghar, Maharashtra

Dr. Ramdas Lihitkar

Librarian, Institute of Science, Nagpur.

Abstract:

The trend of data-fueled research is appearing in all sectors, creating a need for librarians to collaborate with other disciplines to fill a service gap. Libraries and information, are adapting to accommodate for the growth of data by providing data literacy instruction. Academic and research libraries have been through an evolution; This paper started with the introduction and summarizes the opportunities and challenges with big data in Libraries.

Keywords: Big Data, Challenges, Issues, library

Introduction:

Libraries play an important role at the intersections of government, research institutes, and the public since they are storing and managing a large amount of data and those data in library need to be transformed into knowledge which then be used by researchers or users. Librarians might not know how to transform, analyze, and present data in order to facilitate knowledge. For example, they should know how to make big datasets more useful, visible. With new and powerful analytics of big data.

“Big data” describes innovative techniques and technologies to store, distribute, manage and analyze datasets that traditional data management tools can't handle. The concept of “Big data” was first defined by Lance F. Davis. According to the definition, big data is mainly characterized by three V's: Volume, Velocity,

refers to the data volume. General speaking, the size of the data sets compared to regular data. However, it seems that there is no fixed number, how big of data could be classified as big data. Therefore, the size of the disciplines. Traditional software usually can handle megabyte and gigabyte sized data sets, while big data tools should be able to handle terabyte sized data sets. Velocity refers to the situation where data is created dynamically and fast. The speed could be second or so. The third V, refers to variety, which makes big data sets hard to analyze. The regular type of data collected by researchers or business is structured, such as data entered into a spreadsheet with specific rows and columns. Big data might have unstructured data and different types of data, such as

International Data Corporation (IDC)

Technologies describe a new generation of technologies and architecture that "automatically extract value from very large volumes of a wild variety of data for capture, discovery and/ or analysis" (2011)

describes big data as "high-volume, velocity and/or variety information assets that require innovative, innovative forms of information processing that enable enhanced analytics and process automation." This ties to Weinberger's concept of the data can be made smart and piled together to meet any specific users need. He says they are not like catalog cards with more room and an extra fourth IQ points. He says it is like an ISBN that enables distributed information to come together when needed to lead into a genius."

Libraries and Librarians

Information professionals, such as Information Manager, Network manager, Librarian, facilitator/Trainer/ Innovator and also Frontier of latest technology and innovation, have an important role to fill in the future of big data. Librarians have the mindset, skills and ability to help large organizations of all kinds to make the best use of these massive

Management - facilitating use of external datasets

- Data literacy – helping users exploit existing data resources + training faculty about managing research data – preparing graduates for jobs in employment
- Collection building – auditing and appraising data assets + creating metadata/cataloging selected dataset
- Digital curation – capturing, organizing, preserving, and archiving data generated by students and faculty
- Publishing support – advising researchers on identifying, citing, and demonstrating impact of data
- Policy development – consulting stakeholders, drafting and advocating of take-up and implementation

Emergent Roles of Libraries and Librarians

- Data literacy – preparing frontline library staff to respond to public digital privacy or data profiling + training youth librarians to support informal learning programs for teens in public libraries
- Metadata consultancy – providing specialized expertise to support data by municipal government agencies
- Infrastructure development – serving as local platforms for open developing data in smaller communities (e.g., hosting data + facilitating data deposit) + building, testing, and evaluating custom storing, transferring, and processing Big Data for re-use
- Data protection – promoting responsible use of personal data

Characteristics of Big Data:



These characterizations represent ten different challenges associated with the handling of data (as capture, cleaning, curation, integration, storage, processing, archiving, transfer, mining, analysis, and visualization).

Opportunities & Challenges for Libraries

Libraries have a long history of collecting data and reporting their analyses. Library data collection focused on gathering information about library materials, library service activities. The data were often compiled into library reports as a way to assess a library's resources and performance. In recent years, assessment has grown significantly in the area of assessment as a way to demonstrate and accountability to various stakeholders. Academic libraries have been an important and leading part in this movement as well. The libraries have developed new assessment tools and methods and expanded our data collection to include quantitative data (interviews, chat transcripts, etc.), social engagement data (Twitter, etc.), usability testing, and collection analysis, just to name a few (Cox & Jantti, 2012) and Research Libraries, 2010). Furthermore, the rise of Big Data has made data collection tasks easier and faster; it also has enabled libraries to move beyond simple counting and compiling statistical measures and to engage in complex data analysis, data mining analytics (Cox & Jantti, 2012) and research performance analysis

Libraries and data centers opportunities
Lower barriers to researches to make their data available
Integrate data sets into retrieval services
Support of persistent identifiers
Engage in developing common meta-description schemas and common citation practices
Promote use of common standards and tools among researchers
Support crosslink's between publication and datasets
Provide and help researchers understand meta-description of datasets
Establish and maintain a knowledge base about data and their context
Curate and Preserve datasets
Archive software needed for re-analysis of data
Be transparent about conditions under which data sets can be re-used (expert knowledge needed, software needed)

	Engage in establishing uniform data citation standards
Citability	Support and promote persistent identifiers Transparency about curation of submitted data Promote good data management practice
Curation /Preservation	Collaborate with data creators Instruct researchers on discipline specific best practices in creation (preservation formats, documentation experiment)

Issues with Big Data in Library

There are some issues which are common to library big data research as listed below.

A. Lacking of Data Scientists

According to studies, USA might not be able to fill half of the positions of data scientists and data managers by 2018 . The situation in library might be same. The key reason is that data analysts need not only the skills of statistics and computer science, but also a lot of domain knowledge and collaboration ability. Therefore, the challenges faced by library are the ability to manage the information of big data. It seems that short-course training will not be sufficient.

B. Ability of Adopting Big Data

Big data comes in various fields. However, a lot of companies are not ready to adopt it. According to the study, more than half of organizations could not handle the big data currently due to lack of personnel and platform .Research of library big data is even slower than that in other disciplines. The key reason is that the digital libraries tend to be contained organizational units and they try to stay back from new technology.

C. Budget Issues

Although more and more people understand the great benefit of using big data analysis, the IT investment such as analytics servers, high-performance computing servers is needed. Majority of US government organizations have not had plan for investment in big data mainly due to budgetary issue . It seems that most of library administrations have not placed big data on the table because of shrinking budgets as well Research data management projects are paid less attention due to the challenge of human resources. Moreover, a lot of research data which were produced ten year ago is still analogue, such as the library

for biology research or geology work. Digitizing these resources is not a simple job and need a lot of time and personnel resource.

D. Technical Challenges

Big data involves techniques such as capturing, storing, processing and presenting data. Data in the library have different types and might be in various statuses. Some data might be difficult for digitalization. For geological data, data capturing often face challenge. For biological data, digitalizing field trip notes and geological maps is still an issue. On the other hand, some of data often contains some dirty or false data. Therefore, correctly removing those data need some work. Due to heterogeneous types and formats of research data, integrating data become a very tough job. For example, the challenge of integrating earth science data is not only in geological library, since data across multiple disciplines (geology, geography, geology and hydrology) has been collected, managed, and documented in very different formats. Many types of research data are considerably less usable when they are in their raw form than after they have had filters or algorithms or other processing performed on them. Therefore, work need budget to build tools and provide other supports as well.

E. Privacy

Big data is mining the data and discovering knowledge. There should be a privacy protection. On the other hand, new risks of system intrusions might arise due to the accessibility of big data and amount of data. Data security issues have not been well considered for library big data research.

F. Big Data Not for all Organizations

It is clear that the organizations that plan to use big data need to have a relative large investment in IT infrastructure and personnel. Therefore, small library without enough budget support might need to share the resource with other organizations. On the other hand, big data is a relative new and traditional analytic approach still dominates majority of organizations. In regards to the individual research data, small library might not have enough resource to support direct interaction with research faculty. Therefore, it might be hard to integrate all the data from all researchers in the organization.

Conclusion:

Big data has its own importance in the world of technology. Day by day big data is used and generated by libraries. It is largely the librarians' role to create and

provide data management services that meet this new grant specification. The data entered into the data lifecycle by taking on the task of housing and managing the data generated from student and faculty research projects in data repositories.

References:

1. R. H. (2015). The Next Frontier federal Librarians and Data . IBM
2. Alex, H. (2012, June 26). <http://strata.oreilly.com/2012/06/prepare-big-data-nvc>. Retrieved Jan 24, 2018, from strata oreilly.com.
3. <http://1j.libraryjournal.com/2013/01/oa/what-governmental-big-data-means-for-libraries>. (2013, May 30). Retrieved Jan 24, 2018, from libraryjournal.com.
4. Kent .A. L. .. (1986). Encyclopedia of Library and Information science. Marcel Dekker.
5. Lehong.H. L. .. (2013, March 1). Tooikit Board ready Silder on Big Data Opportunities Gartner.
6. President. E. o. (2013). Open Data Policy Managing. Washington
7. Schwartz, M. (2013, May 30). <http://1j.libraryjournal.com/2013/01/oa/what-governmental-big-data-may-mean-for-libraries>. Retrieved Jan 24, 2018, from libraryjournal.com.
8. Weinberger.D. (2007). Everything is Micellaneous The Power of Organizing. New York: Times Books.
9. Exposing library data with big data technology: A review (PDF) Available https://www.researchgate.net/publication/306926515_Exposing_library_data_with_big_data_technology_A_review [accessed Jan 24 2018].
10. Cox: B., & Janiti, M. (2012). Discovering the impact of library performance. EDUCAUSE Review Online. Retrieved February 2018. <http://www.educause.edu/ero/article/discovering-the-impact-of-library-performance>