

Robotic Technology, Artificial Intelligence and the 4IR Libraries

Current Topics in Library and Information Practice

Robotic Technology, Artificial Intelligence and the 4IR Libraries

Edited by

Adeyinka Tella and Omwoyo Bosire Onyancha

e-ISBN (PDF) 978-3-11-117698-7 e-ISBN (EPUB) 978-3-11-117754-0 ISSN
2191-2742

Library of Congress Control Number: 2024944104

Bibliographic information published by the Deutsche Nationalbibliothek The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available on the internet at <http://dnb.dnb.de>.

© 2025 Walter de Gruyter GmbH, Berlin/Boston, Genthiner Straße 13, 10785 Berlin Typesetting: Integra Software Services Pvt. Ltd.

Printing and binding: CPI books GmbH, Leck

www.degruyter.com

Questions about General Product Safety Regulation:
productsafety@degruyterbrill.com

Contents

The Authors IX

Lorette Jacobs

Chapter 1

Internet of Things (IoT)-based Smart Libraries: Opportunities and Advances
1

Amos Orenyi Bajeh, Adeyinka Tella

Chapter 2

Internet of (Robotic) Things in Library System Automation 21

Tom Kwanya

Chapter 3

Enhancing the Findability of Library Information Resources Using the Internet of Things: A Kenyan Perspective 53

Olaronke O. Fagbola

Chapter 4

The Prospect of Blockchain Technology Adoption in Academic Libraries: An Overview 73

Rhodes Elias Mwageni and Robert Stalone Buwule

Chapter 5

The Assumptions of the Use of Artificial Intelligence Technologies in Robotics in Inclusively Engaged Academic Libraries in Sub Saharan Africa: A Literature Survey 95

Airen Edale Adetimirin and Oyedele Abimbola Ojeniyi

Chapter 6

Building Future Information Professionals for 4IR Libraries: Role of Library and Information Science Lecturers in Nigeria 111

Temidayo Oluwatosin Omotehinwa

Chapter 7

Artificial Intelligence and Libraries: The Possibilities in the 4IR 131

VI Contents

Robert Stalone Buwule and Stephen Mutula

Chapter 8

Inclination of Uganda University Libraries on Co-Driving the Fourth Industrial Revolution 145

Adeyinka Tella, Joseph. M. Ngoaketsi and Oluwatoyin Oluwayemi Obinyan

Chapter 9

Legal and Ethical Issues Associated with Adoption and Utilisation of Cloud Computing in Academic Libraries in Nigeria 165

Wilson Chukwunedum Ochonogor and Edwin Iroroeavwo Achugbue

Chapter 10

The Relevance of Cloud Computing and the Internet of Things for Effective Information Service Delivery in Medical Libraries in Nigeria 197

Aderinola Ololade Dunmade, Uloma Doris Onuoha, and Adeyinka Tella

Chapter 11

Cyberethics and Cybersecurity in Universities in Developing Countries: A Literature Survey and Course Content Analysis from Library Instruction's Perspective 215

Musediq Tunji Bashorun, Kabir Muhammad Abubakar, Iro Sani Lawal, and Rafiat Bukola Bashorun

Chapter 12

Application of Big Data and Advanced Analytics in Future Librarianship 227

John Jackson Iwata

Chapter 13

Libraries in the Wake of the Fourth Industrial Revolution: A Portrait of Roles, Challenges, and the Prospects 249

Collence Takaingenhamo Chisita, Rexwhite Tega Enakrire, and Blessing Chiparausha

Chapter 14

The Changing Role of Librarians in the Age of Hypermedia Seduction: Reflection on the Researchers' Exposition 267

Contents VII

Dayo Reuben Aremu

Chapter 15

Machine Learning and Data Mining: An Overview for Emerging Trend in Fourth Industrial Revolution 287

Yusuf Ayideji Ajani, Ganiyu Ojo Adigun, Saheed Oyeniran, and Olufunke Abosede Ayeni

Chapter 16

The Fourth Industrial Revolution and the Relevance of its Technologies to Libraries 309

Adeyinka Tella

Chapter 17

Artificial Intelligence Humanoid Robots for Public Libraries Reference Services 343

Omwoyo Bosire Onyancha

Chapter 18

Data Science: A Systematic Review of the Concept Using Author-supplied Keywords 357

Petros Nhlavu Dlamini, Yusuf Ayodeji Ajani, and Ugonna Vivian Ailakhu

Chapter 19

The Prevalence of Digital Pedagogy for Delivering Library Instructions in the Era of the Fourth Industrial Revolution 379

Abdulumumin Isah, Mariam Oyinda Aborisade, and Stephen Mayowa Omolounnu

Chapter 20

Awareness and Usage Readiness of Artificial Intelligence in University Libraries in Kwara State, Nigeria 403

Adeyinka Tella and Ajani Yusuf Ayodeji

Chapter 21

Virtual/Augmented Reality Meeting Platforms: The Enablers of Librarians' Collaboration in 3D Virtual Environment 421

Index 447

The Authors

Kabir Muhammad Abubakar, Department of Library and Information Science, Faculty of Education, Umaru Musa Yar'adua University Katsina, Katsina State, Nigeria.

muhammad.kabir@umyu.edu.ng

Edwin Iroroeavwo Achugbue, Department of Library & Information Science, Faculty of Education, Delta State University, Abraka.

edwin12ng@gmail.com

Airen Edale Adetimirin, Department of Library, Archival and Information Studies, University of Ibadan, Ibadan, Nigeria.

aeadetimirin@gmail.com

Ganiyu Ojo Adigun, Ladoke Akintola University of Technology, Ogbomoso, Oyo State, Nigeria. adiganfolly@yahoo.com

Ugonna Vivian Ailakhu, Department of Library and Information Science Faculty of Education National Open University of Nigeria, Abuja, Nigeria.

vailakhu@noun.edu.ng

Yusuf Ayideji Ajani, Graduate Students Department of Library and Information Science, University of Ilorin, Nigeria.

trustusouph@gmail.com

Dayo Reuben Aremu, Department of Computer Science, Faculty of Communication, University of Ilorin, Ilorin.

Olufunke Abosede Ayeni, Dept of Library and Information Science Federal Polytechnic, Offa. Kwara State, Nigeria.

bosfunke@gmail.com

Ajani Yusuf Ayodeji, Department of Library and Information Science, University of Ilorin, Ilorin, Nigeria.

trustusouph@gmail.com

Amos Orenyi Bajeh, Department of Computer Science, University of Ilorin, Ilorin, Nigeria. bajehamos@unilorin.edu.ng

Musediq Tunji Bashorun, Department of Library and Information Science, Faculty of Education, Umaru Musa Yar'adua University Katsina, Katsina State, Nigeria. musediq.bashorun@umyu.edu.ng

Rafiat Bukola Bashorun, The Polytechnic, Ibadan, Nigeria.
rafiatbashorun@gmail.com

Robert Stalone Buwule, University Librarian, Mbarara University of Science and Technology. Dr. Robert Stalone Buwule is a University Librarian of Mbarara University of Science and

Technology. Previously he was a Senior Librarian at Kyambogo University (KyU) where he headed the Library ICT section. He obtained his doctorate from the University of KwaZulu Natal in South Africa. He holds a Master's of Information Science from Makerere University in Kampala, Uganda, and a

<https://doi.org/10.1515/9783111176987-203>

X The Authors

Master's of Information Technology at the University of Pretoria, South Africa. His research interests focus on the application of technological innovation, improving information access for SMEs, persons with disabilities, and empowerment of youth and other marginalized persons through information products and services. He has presented research papers at the Standing Conference of East, Central, and Southern African Librarians (SCECSAL), IFLA WLIC satellite conferences, IS and PROLISSA conferences.

rbuwule@must.ac.ug or burosta@gmail.com

Blessing Chiparausha, Bindura University of Science Education, Zimbabwe.

Collence Takaingenhamo Chisita, University of South Africa, Pretoria.

Petros Nhlovu Dlamini, Department of Information Studies, Faculty of Arts, University of Zululand. DlaminiP@unizulu.ac.za

Aderinola Ololade Dunmade, Centre for Open and Distance Learning, University of Ilorin, PMB 1515, Ilorin, Nigeria.

Rexwhite Tega Enakrire, University of Johannesburg, South Africa.

Olaronke O. Fagbola, University of Ilorin, Nigeria.

Dr. Olaronke O. Fagbola, Ph.D., CLN, is a lecturer in the Department of Library and Information Science, University of Ilorin, Nigeria. She holds a doctorate in Library and Information Studies, Master degrees in Information Science and Master in Business Administration from the prestigious Universities of Ibadan and Lagos respectively and a Bachelor of Science degree in Education/ Biology from the Obafemi Awolowo University, Ile-Ife. Her research focuses on information management, users' studies, bibliometrics, and technology in academic libraries.

faith1931@yahoo.com or fagbola.oo@unilorin.edu.ng.

Abdulumumin Isah, Department of Library and Information Science, University of Ilorin.

John Jackson Iwata, Department of Knowledge Management, Moshi Co-operative University, Moshi-Kilimanjaro, Tanzania.

John Jackson Iwata is a teacher and library expert. He holds a Ph.D. in Information Studies of the University of KwaZulu-Natal, and both Master of Arts in Information Studies and Bachelor of Education in Arts of the University of Dar es Salaam. Besides lecturing and managing library collection, from 2010 to 2020 Mr. Iwata was involved in the management team of Moshi Co-operative University as Head of Periodicals Department, Deputy Director, and the Director of Co-operative Library and Archives. He was the first chairperson of the Steering Committee of establishing Cooperative Archives Centre at the same University in 2010. Also, he has been involved as a team leader in designing training programs in library and information sciences, and in developing some university guidelines and policies. Mr. Iwata is a former Postdoctoral Research Fellow of the University of KwaZulu-Natal, South Africa. During his postdoctoral studies, he assisted in the University Curriculum Transformation Project (UCDP) specifically on the Decolonization and Transformation of University's Education Curriculum in South Africa. His research areas include knowledge management (specifically management of indigenous knowledge), information literacy, and curriculum development.

iwata2j@gmail.com ORCID: <http://orcid.org/0000-0002-4491-650X>

Lorette Jacobs, Department of Information Science, University of South Africa, Pretoria, South Africa.

ejacobl1@unisa.ac.za

Tom Kwanya, The Technical University of Kenya. tkwanya@tukenya.ac.ke

Iro Sani Lawal, Department of Library and Information Science, Faculty of Education, Umaru Musa Yar'adua University Katsina, Katsina State, Nigeria.

sani.lawal@umyu.edu.ng

Stephen Mayowa Omolounnu, Department of Library and Information Science, University of Ilorin

Stephen Mutula, University of Kwazulu Natal.

Rhodes Elias Mwageni, Director of Library Services, Kilimanjaro Christian Medical University College (KCMUCo).

(A Constituent College of Tumaini University Makumira)

Dr. Rhodes Elias Mwageni is a Director of Library Services at Kilimanjaro Christian Medical University College (KCMUCo) (a Constituent College of Tumaini University Makumira (TUMa) in Tanzania in East Africa. He completed his BA General (Hons) degree, Postgraduate Diploma in Education (PGDE), and MA in Information Studies at University of Dar es Salaam (in Tanzania), and his Ph.D. (Information studies) at the University of KwaZulu Natal, Pietermaritzburg, in South Africa. His research interests include embracing reading culture as an attitude and behaviour, organizational behaviour, education, marketing of information services, reference services, digital library services, artificial intelligence technologies in robotics, and extensive experiences of research supervision at different levels.

rhodes.mwageni@kcmuco.ac.tz or rhodesmwageni@gmail.com

Joseph. M. Ngoaketsi, Department of Information Science, University of South Africa. ngoakmj@unisa.ac.za

Wilson Chukwunedum Ochonogor, Department of Library & Information Science, Faculty of Education, Delta State University, Abraka.

wilsonchukwunedum@gmail.com

Oyedele Abimbola Ojeniyi, Central Bank of Nigeria-Centre for Economics and Finance, University of Ibadan, Ibadan, Nigeria.

deleojeniyi@yahoo.com

Oluwatoyin Oyeyemi Obinyan, The Library, Edo State University, Ekpoma, Nigeria.

Oluwatoyin Obinyan is the acting Head of the Electronic Library Department of Ambrose Alli University Library, Ekpoma Edo State Nigeria. She is a doctoral student at the University of Nigeria, Nsukka, with her areas of specialization including Information Literacy skills, Electronic resources management and use, ICT application to library processes, and Academic librarianship. obinyantoyin@gmail.com or obinyantoyin@aauekpoma.edu.ng

T. O. Omotehinwa, Department of Mathematics and Computer Science, Federal University of Health Sciences, Otuokpo, P.M.B. 145, Nigeria.

Uloma Doris Onuoha, University Library, Adeleke University, Ede, Osun State, Nigeria.

XII The Authors

Omwoyo Bosire Onyancha, Department of Information Science, University of South Africa, PO Box 392 Unisa, 0003.

onyanob@unisa.ac.za

Saheed Oyeniran, Department of Educational Management, University of Ilorin, Nigeria. saheed.oyeniran@gmail.com

Mariam Oyinda Aborisade, Department of Library and Information Science, Osun State University.

Adeyinka Tella, Department of Library and Information Science, University of Ilorin, Ilorin, Nigeria; Department of Information Science, University of South Africa, Pretoria.

Adeyinka Tella is a Professor and Head of Department at the Department of Library and Information Science at the University of Ilorin in Nigeria and is currently a DISTINGUISHED PROFESSOR at the University of South Africa in

Pretoria, South Africa. tellayinkaedu@yahoo.com or tellaa@unisa.ac.za or tella.a@unilorin.edu.ng

Musediq Tunji Bashorun, Kabir Muhammad Abubakar,
Iro Sani Lawal, and Rafiat Bukola Bashorun

Chapter 12

Application of Big Data and Advanced Analytics in Future Librarianship

Abstract: Big Data is increasingly getting more attention in all sectors of organizations; in business, academia, and research including libraries. Big Data is a by-product of the information explosion. Due to advancements in technologies, modern life produces data at an amazing rate which lead to advances in data storage and analysis and the arrival of the concept of Big Data. The chapter briefly describes the concept of big data, the characteristics of big data, and the history of Big Data. Furthermore, the chapter examines the differences between traditional data and Big Data. It also covers Big Data technologies and the relevance of Big Data to all sectors. It covers Big Data analytics and applications of big data in the libraries. Also, it highlights the benefits of Big Data in various sectors including libraries. The chapter covers Big Data analysis techniques. Moreover, the chapter identifies challenges facing big data and advanced data analytics. It is concluded that Big Data technology has huge benefits to various sectors, especially libraries, and its adoption has come to stay for the effective and efficient utilization of information for the benefit of information users.

Keywords: Analytics, Big Data, Libraries, Traditional Data, Users

Introduction

The era of information explosion is now, with Big Data today becoming a catchphrase in every area of the global economy. The digital era of information society evolved with data, information, and knowledge creation. Its information explosion stage transformed the world information organization into data mining, with meta-data, semantic web, webometrics, ontology, and so on (Witt et al., 2018). However, Big Data is a term for a “collection of data sets so huge and complex that it becomes difficult to process using on-hand database management tools or traditional data processing applications” (Oguntimilehin & Ademola, 2014, p. 433). There is no univer-

sally recognized operational definition of Big Data According to Gordon-Murnane (2012), Big Data is not just about gigabytes of data but immeasurably increased. International Business Machines (IBM) reported that “2.5 quintillion bytes of data is

<https://doi.org/10.1515/9783111176987-012>

created on a daily basis.” Therefore, Big Data could be defined as the new technologies and architectures designed contain huge amounts of structured, semi-structured, and unstructured types of data and are difficult to process using traditional databases and other software.

Commercial and Industrial Security Corporation (CISCO) (2012) predicts that, between 2015 and 2016, global IP traffic will grow to more than 330 Exabytes. It can be deduced that Big Data is loosely structured data that is often incomplete and inaccessible which may consist of billions to trillions of records of millions of people, all from different sources, for example the Web, sales, customer contact centers, social media, mobile data, and so on. That is almost more data than the two previous years combined. Gordon-Murnane (2012, p. 30) reported that McKinsey Global Institute’s statement indicated that one Exabyte of data is the equivalent of more than 4,000 times the information stored in the US Library of Congress. The starting of terabytes may be considered as the start of Big Data.

These show how big Big Data is. Big Data has become a major issue of concern in all types of organizations whether business or service, private sector or public sector, including libraries and information centers. It can be inferred that the Big Data effect is being felt everywhere, from business to science, from the government to the arts as supported by Cukier (2010). Due to the “arrival of the Big Data concept, decisions that are previously based on guesswork or painstakingly constructed models of reality can now be made based on data itself” (Oguntimilehin & Ademola, 2014, p. 434). The exponential increase in Big Data continuously challenges the storage industry. However, physical storage could be enhanced with more terabyte drive arrays and the software infrastructure could be made more flexible to quickly and economically accommodate greater volumes of transactions and queries (Pan et al., 2018). Some outstanding achievements of Big Data are the Los Angeles Police Department and the University of California using Big Data to predict crime before it even happens and Google Flu trends using search terms to predict the spread of the flu virus. Also, Statisticians Nate Silver predicted the outcome of the US election down to each individual state in 2012, and MIT is using mobile phone data to establish how people’s location and traffic patterns can be used for urban planning (Oguntimilehin & Ademola 2014, p. 435).

Big data analytics is the process of breaking down data sets into smaller chunks containing a variety of data types for efficient processing to uncover hidden patterns, market trends, customer preferences, unknown co-relations, and useful business information (Bhadani et al., 2016). Big Data analysis aims to simplify the character and scope of growing knowledge in different fields of science;

especially multidisciplinary signal processing approaches (Slavakis, Kim., Mateos & Giannakis, 2014).

Big Data is a big deal and it is a huge opportunity for librarians and information professionals to play a key role in the big data universe. Librarianship is a profession concerned with acquiring, organizing, and disseminating collections of books and related materials in libraries and servicing readers and others with these resources. These activities have to be carried out by an expert referred to as a librarian. A librarian is a person who works professionally in a library, providing access to information, social, and technical programming required by users to satisfy their information needs. The future starts from now, and librarians must abide by the current trend of information explosion by adopting the effective utilization of huge data tagged Big Data. Therefore, the paper describes the application of big data and advanced analytics in future librarianship.

History of Big Data

The term data was first used in 1941 in the Oxford English Dictionary. Thereafter, researchers and librarians frequently began to write on information and data explosions. Fremont Rider wrote a book in 1944 through Wesleyan University Librarian titled *"The scholar and the future of the research library"*. The author estimated that American universities were doubling in size every 16 years. Based on this growth rate, Rider had speculated that the Yale Library in 2040 would have "approximately 200,000,000 volumes which will occupy over 6,000 miles of shelf and require a cataloging staff of over six thousand persons (Bhoola et al., 2014). In 1961 Derek Price published *Science since Babylon*, in which he charted the growth of scientific knowledge by looking at the increased numbers of scientific journals. The papers were shown to grow exponentially, doubling every 15 years and increasing by a factor of ten over half a century (Alvarado, 2016). Derek Price calls this the "law of exponential increase." This assertion suggests that there is a rapid increase in the number of scientific journals that leads to exponential growth.

In July 1986, Becker estimated that "the recording density achieved by Gutenberg was approximately 500 symbols (characters) per cubic inch – 500 times the density of (4,000 B.C. Sumerian) clay tablets." In 1996, storage systems had evolved to meet the changing customer needs and digital storage became more popular. This leads to the development of autonomous storage (Morris & Truskowski, 2003). In October 2000, Lyman and Varian carried out the first comprehensive study at the University of California to quantify the total amount of new and original infor-

mation in computer storage terms and established that the world produced between one and two Exabytes of unique information per year. This is roughly 250 megabytes for every man, woman, and child on earth and digital content was dominant.

In February 2001, Doug Laney, an analyst with the Meta Group, published a research note entitled “3D Data Management Controlling Data Volume, Velocity, and Variety”. A decade later, the “3Vs” has generally been accepted as the three defining dimensions of Big Data (Oussous et al., 2018; Koga, 2015). In September 2005, Tim O’Reilly published “What is Web 2.0” and he asserted that “data is the next Intel inside.” In January 2008, Bret Swanson and George Gilder published “Estimating the Exaflood (PDF)” and projected that U.S IP traffic could reach one Zettabyte by 2015 and that the US Internet of 2015 would be at least 50 times larger than it was in 2006. Kenneth Cukier in February 2010 published an Economist Special Report titled “Data, Data everywhere” and wrote, “the world contains an unimaginably vast amount of digital information which is getting ever vaster more rapidly (Anderson & Rainie, 2018).

In May 2011, Manyika et al. (2011) carried out a study and estimated that in 2010 7.4 Exabytes of new data were stored by enterprises and 6.8 Exabytes of new data were stored by consumers (on their PCs and other devices). It can be deduced that due to the widespread accessibility, affordability, and availability of new digital devices and easy access to the Internet, there has been a data explosion. Every day, we create 2.5 quintillion bytes of data, so much that 90% of the data in the world today has been created in the last two years alone. The data comes from everywhere sensors used to gather climate information, posts on social media sites, digital pictures, and videos. Also, in purchasing transaction records and cell phone GPS signals, to name a few, such data is Big Data (IBM Newsletter, 2012). Social media is generating petabytes of data per day. Syed (2015) reported that Joe Hellerstein, a computer scientist at the University of California in Berkeley, calls this age the industrial revolution of data. This is the stage of the arrival of Big Data with its various characteristics.

Characteristics of Big Data

The term Big Data covers every piece of data generated from organizations that have been stored until now. This includes data stored in clouds and even in the URLs that you bookmark. According to Olendorf and Wang (2017, p. 192), Big Data refers to any data that exceeds the user’s ability to analyze it in one of three dimensions (Volume, Velocity, and Variety). Big Data is characterized by the three

Vs' high volume, high velocity, and high variety (Sonawane, 2018; Bieraugel, 2015). Furthermore, Lomotey and Deters (2014) extended the model defined by Laney (2001) to a 5Vs model (volume, veracity, velocity, value, and variety).

Moreover, several studies (Akoka, Comyn-Wattiau, & Laoufi, 2017; Khan et al., 2017) were carried out and the findings revealed more features that extended the model from a 5Vs to a 6Vs model. Big Data is now characterized by 6Vs (volume, velocity, variety, value, viability, and veracity). The six (6) characteristics of Big Data are shown in Figure 1.

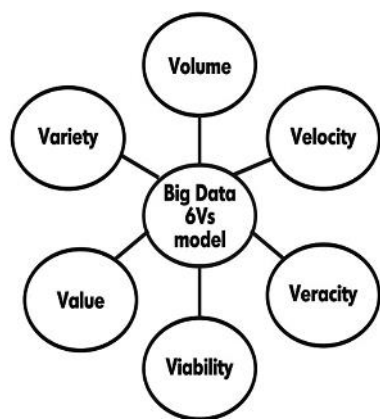


Figure 1: The 6Vs model that currently defines Big Data.

Source: adapted from (Ristevski & Chen, 2018).

Volume: Big Data is huge; it is machine-generated data produced in much larger quantities than non-traditional data depending on industry or discipline. This data is loosely defined data that cannot be stored or analyzed by conventional hardware and software. Traditional software can handle megabyte and kilobyte-sized data sets, while big data tools can handle terabyte and petabyte-sized data sets. This is echoed by Manyika et al. (2011), who assert that Big Data refers to datasets whose size is beyond the ability of typical database software tools to capture, store, manage, and analyze (Bieraugel, 2015). With the application of the Internet, there is an enormous growth of information leading to information overload (explosion). This has led to an ever-increasing userbase and increased demand for a huge amount of data that requires extra attention.

Velocity: This is data created from a wide array of data platforms: for an instant, the speed at which someone can create a single tweet on Twitter, or post to Facebook (Bieraugel, 2015). Velocity could be described as the speed at which data is processed and becomes accessible to meet demands and challenges that lie in the path of development and growth. It is generally created by social media platforms

where a large influx of opinion and relationships management records is created. It is high frequency, for example Twitter data, posts and likes on Facebook, and many others. For time-sensitive processes such as detecting fraud, Big Data must be used as it streams into enterprises to maximize its value (IBM Newsletter, 2012). The same velocity issue prevails in any form of the library by creating a large volume of information from large collections at high speed. Increasingly modern libraries are using social media platforms and other digital web-based information resources such as search engines and subject portals, institutional repositories, and open access journals which generate large amounts of data. Typically there is good information hidden amongst a larger body of non-traditional data; the challenge is to identify what is valuable and then transform and extract that data for analysis.

Variety: This could be described as all structured and unstructured data that has the possibility of being generated, either by humans or by machines. It is all about the ability to group the incoming data into different categories. In addition to the velocities and varieties of data, data flows can be highly inconsistent with periodic peaks. Data comes in all types of formats from structured, numeric data in traditional databases to unstructured text documents: email, video, audio, stock ticker data, and financial transactions. Variety makes big data sets more challenging to organize and analyze. Big Data sets can contain unstructured data such as email messages, photographs, postings on internet forums, and even phone transcripts (Bieraugel, 2015). In the twenty-first century, libraries are overwhelmed with a variety of data sets, both structured and unstructured (De Mauro, Greco, & Grimaldi, 2016). Structured data may include e-books, institutional repositories, open access journals, and audio-video while unstructured data may contain emails, postings on social media platforms, metadata, and others (Power, 2014). This variety of unstructured data creates problems for storage, mining, and analyzing of data (Normandeau, 2013). There is a diversity of information retrieved from various sources with diverse forms and formats.

Value: This refers to the usefulness of gathered data for a specific task. It can also be described as the value that Big Data can provide to any organization. According to Cartledge (2016), it is the source of Big Data, whether it is reliable, accurate, and complete. It enables and empowers users to be able to convert information into something valuable and useful.

Viability: This can be described as a careful selection of those attributes in the data that are most likely to predict outcomes that are germane to organization. With so many varieties in data and variables to consider, it is important to assess the viability of data to build an effective predictable model (Biehn, 2013) and also to expand the model.

Veracity: This has to do with how accurate and truthful a data set may be. Big Data generates a lot of data and a variety of data. However, the data must have quality and produce credible results that enable action to produce the right decision-making. The veracity determines the accuracy of the data (Leonelli & Tem-pini, 2021; Scott, 2015; Cai & Yangyong, 2015). The assertions suggest that the veracity of Big Data is not just the quality of the data itself but the reliability of the data source, type, and processing. In library settings, the information generated must be authentic, credible, and accurate information that reduces inconsistencies and uncertainty.

Today's data comes from multiple sources, which makes it difficult to link, match, cleanse, and transform data across systems. However, it's necessary to connect and correlate relationships, hierarchies, and multiple data linkages or else your data can quickly spiral out of control. IBM believes that Big Data exploration, an enhanced 360-degree view of the customer, security/intelligence extension, operational analysis, and data house augmentation are the ways to get started with big data (Jewell et al., 2014). The following steps are for the handling of Big Data:

1. Understand the Big Data sources for your business, i.e. streaming data (data generated by your IT system from the web), social media data, and publicly available sources.
2. Query/ consider the decisions you need for analysis. It will include how to store and manage the data, how much of it to analyze, and how to form strategies or insights that are uncovered or covered.
3. Do research for technologies that manage Big Data and its analytics (cheap for abundant storage, faster processor, affordable open-source, distributed big data platforms, high connectivity, parallel processing, cloud computing, and others).

Differences Between Traditional Data and Big Data

Big Data is a huge amount of structured, semi-structured, and unstructured types of data that is so large that it's difficult to process using traditional databases and other software. The traditional database could be described as a centralized database architecture used to store and maintain the data in a fixed format or fields in a file. In addition, traditional data means structured data which is being majorly preserved by all types of businesses starting from very small to big organizations.

There is a need to differentiate between traditional data and Big Data.

S/N	Traditional Data	Big Data
□	Data generated at the enterprise level	Data generated at enterprise-level and outside level
□	Data stored in the form of structured and unstructured	Data is stored in the form of structured, semi-structured, and unstructured
□	Data manage in centralized form	Data manage in a physical distributed form
□	Specialized high-level software as well as the hardware used	Inexpensive commodity boxes in cluster mode
□	Data stored in gigabytes or terabytes	Data stored in Petabytes, Zettabytes, Exabytes
□	By default, it has a stable and interrelationship	Unknown relationship
□	Traditional data sources include ERP transaction data, CRM data, Web transaction data, financial data	Non-traditional data sources include social media, log data, device data, sensor data, video images

In addition, examples of sources where Big Data could be generated are:

1. YouTube users upload 72 hours of video each minute
2. The average numbers of tweets per day is 58 million
3. Every 60 seconds there are 695,000 status updates, 79,364 wall posts and 510,040 comments published on the social networking site Facebook.

All these activities can only be achieved with the help of a set of technologies.

Big Data Technologies

Big Data not only refers to the data itself but also a set of technologies that capture, store, manage, and analyze large and variable collections of data to solve complex problems. Big Data takes too much time and costs too much money to load into traditional relational databases for analysis. Big Data analytics is often associated with cloud computing because it encompasses the analysis of large data sets in real-time where a massive volume of both structured and unstructured data is analyzed using different techniques (Sivarajah et al., 2017)

Big Data analysis is not a typical database management system; it requires high-level skills like data analysis, data warehousing, data transformation, and data collection. Several key components widely make use of Hadoop in deep analytics of Big Data, such as garage subsystem, caching layer, scheduler, data practitioner, fact codes, and compression algorithm (Herodotou, Lim, & Luo, 2011).

Hadoop sub-initiatives, including Hive and HBase, provide additional statistical control for storing unstructured and semi-dependent data units (Begoli & Horey, 2012). Big Data also describes the storage and analysis of large and/or complex datasets using a series of techniques, including NoSQL, MapReduce, and machine learning (Blascheck et al., 2015; Ward & Barker, 2013) as well as many others as shown below:

1. **Hadoop:** This is open-source software that is made to handle big data. Hadoop was developed by Doug Cutting in 2005 and first released in 2011 to support distribution for the Nutch search engine project at Yahoo (Bhathal & Amardeep, 2019). It is a device that has grown quickly to be the gold standard in the industry. Hadoop is the most popular source platform for handling Big Data. It can do large-scale processing and is flexible enough to work with multiple data sources (Azeroual & Fabre, 2021). It has several applications such as data from weather and traffic sensors, machine to machine transactional data (Jiang & Luo, 2022).
2. **MapReduce:** This is a programming model comprised of a map procedure method that filters and performs sorting and reduces methods that perform the summary operation. The key contribution of this method is the scalability and fault tolerance achieved for a variety of applications. The in-depth MapReduce platform is now broadly used for clickstream, web indexing and log evaluation, and positive large-scale fact extraction and device mastering functions (Borkar, Carey, & Li, 2012).
3. **Hive:** It is an open-source data warehouse system used for querying and analyzing large datasets that provides data summarization. It is primarily used in data mining applications which otherwise can take several hours to analyze (Dave, 2013).
4. **PIG:** It was developed at Yahoo and can handle and analyze large data sets, and takes less time than a mapper and reducer program. It has two components: a language and the second one called runtime environment where the programs are run.
5. **Cloud Computing:** Cloud computing and Big Data are an ideal combination as it provides the cost-effective and scalable infrastructure to support big data analytics (Ferkoun, 2014).

In addition, there are other Big Data tools as reiterated by Sengupta (2016, p. 91) which have been published for clarity purposes. They are:

6. **SAP in Memory:** SAP's HANA platform offers several benefits over some of the competition, such as the ability to integrate and analyze huge workloads of data to be analyzed in real-time. This is extremely useful for the developer who is looking for speed to market (Darvazeh et al., 2020)

7. **Cambridge Semantics:** Using the Anzo Software Suite, this open platform helps you to collect, integrate, and analyze Big Data to assist in building the Unified Access solutions (Sengupta, 2016). The software has a data integration machine that streamlines data collection and helps with analytics. The key features include being able to combine data from multiple sources and customized dashboards to make analysis easy.
8. **MongoDB:** This is an open-source documental database that is useful to developers who want to have precise control over the final outputs. MongoDB is also scalable and includes third party log tools such as Edda and Fluentd (Sengupta, 2016; p. 91).
9. **Pentaho:** Pentaho joins data integration and business analytics for visualizing, analyzing, and blending Big Data. The open and embeddable platform comes with extensive analytics capabilities with data mining and predictive analysis. The data integration tools mean that users do not require coding in SQL or writing MapReduce Java functions.
10. **Talend:** Straight away, one of the key benefits of Talend's Open Studio is that it is open-source, which means that improvements will keep on rolling out as the community tweaks the tool (Sengupta, 2016). Its tools include products for developing, testing, and deploying data management and application integration products. Additionally, the company manages the full lifecycle, even across enterprise boundaries.
11. **Tableau:** It is one of the most recognized names in the data visualization sphere but it offers many tools for developers that are supported by an active community. Some of the key characteristics of this software are its in-memory analytics database and advanced query language. API, XML, User Scripts, Python, and JavaScript are all supported and so are several browser extensions.
12. **Splunk:** Splunk specifically specializes in harnessing machine data created from several various sources, such as applications, websites, and sensors. The company also enables developers to write code using any technology platform, language or framework.
13. **Splice Machine:** This is a real-time SQL-on-Hadoop database that can help to derive real-time actionable insights, which is a clear benefit for those who are aiming for quick development (Olabode et al., 2022). It is SQL 99 compliant with standard ANSI SQL and can scale from gigabytes to petabytes. As well as support for .NET, Java, and Python, it also offers support for those written in JavaScript/AngularJS.
14. **MarkLogic:** It is built to deal with heavy data loads and allow users to access it through real-time updates and alerts. It provides geographical data that is

combined with content and location relevance along with data filtering tools. This tool is ideal for those looking at paid content search app development.

15. Google Charts: This free tool comes with various capabilities for visualizing data from a website such as hierarchical treemaps or just simple charts. This tool is easily implemented by embedding JavaScript code on a website and allows you to sort, modify, and filter data as well as the ability to connect to a database or pull data from a website. Offering support for popular languages and with the security of knowing that Google will likely keep on improving its offering, this is a good option for many standard developers.

Drawing from the aforementioned tools, it becomes clear that Big Data has enormous benefits accrued to all sectors, especially within libraries.

Relevance of Big Data

The use of Big Data is becoming crucial and the following are the importance of big data (Mcguire et al., 2012).

1. Big Data allows narrower segmentation of customers and this helps in producing more precisely tailored products and services.
2. Sophisticated analytics are useful to policymakers because they would help to improve decision-making.
3. Future generation products and services can be developed with the help of Big Data.
4. There is a significant amount of data that is yet not captured in digital form or made accessible or searchable through the network. Big Data can unlock the value hidden in these forms of data by making information clear.

Big Data is useful in library settings because librarians acquire, process, and store huge amounts of information to satisfy information users. The librarianship profession must strive to meet the needs of all categories of users by embracing and using Big Data.

Big Data Practices in Librarianship

The catchphrase “Big Data” is coming faster and becoming vibrant in business, the techworld, and higher education. The act of gathering and storing large amounts of information for eventual analysis is old. Librarians have always been

great at information management and organization. Librarians are essential knowledge workers. To continue to be relevant and demonstrate librarianship value, there is a need to adopt and use all opportunities created by Big Data. It is a big opportunity for librarians and information professionals to play a vital role in the Big Data universe as library professionals possess the requisite skills, knowledge, and service mentality to help businesses, governments, universities, and non-profits organizations with their information needs.

Nowadays, libraries are offering online resources and services to the credit of Big Data. The libraries are also using social media outlets such as Facebook, Instagram, WhatsApp, and others to promote their services and programs. Librarians with the help of emerging technology and tools such as analytics software can collect more online data, analyze them, and add value to their services. Thus, libraries can use Big Data to make better decisions regarding collection developments, updating information on Current Awareness Services (CAS), updating public spaces, and tracking the use of library materials (Bieraugel, 2013). Librarians can be involved in Big Data by developing skills that can enable data discovery and data retrieval, maintaining data through authentication, archiving, management, preservation, and representation. Librarians can make the big sets more useful, visible, and accessible by creating taxonomies which could be referred to as Big Data.

Application of Big Data and Advanced Analytics in Libraries

One might be thinking that Big Data is not useful to libraries. The library may not collect user data the same way that typical Big Data commercial entities like Amazon or Google or Facebook do. It could be as a result of privacy issues or maybe it is not data needed for the library operations. However, there is often an ocean of data readily available to libraries in various formats. This can be derived from direct observations of library users, cataloguing activities, users' data, survey reports, usage statistics, website metrics, and WorldCat. All these are growing collections of data about library resources, activities, locations, and entities. According to Teets and Goldner (2013), an automation of library circulation systems and the development of WorldCat by OCLC can be called the early applications of big data in libraries.

In the 1950s, decades before anyone pronounced the term "Big Data," businesses were using basic analytics (essential numbers in a spreadsheet that were manually examined or analyzed) to uncover insights and trends. The new benefits that big data analytics bring to the table, however, are speed and efficiency in

data manipulation. Whereas a few years ago businesses would have gathered information, run analytics, and extracted information that could be used for future decisions manually, with Big Data business can identify insights for instant decisions. The ability to work faster and stay agile gives organizations a competitive edge they didn't have before.

Big data analytics assists organizations to connect their data and use it to identify new opportunities. This, in turn, leads to smarter business moves, more efficient operations, higher profits, happier customers, and other information users. The following are the value of Big Data analytics:

1. **Cost reduction:** big data technologies such as Hadoop and cloud-based analytics can be used to identify more efficient ways of rendering library services to users. Also, it brings significant cost benefits when it comes to storing huge amounts of information for the benefit of information users.
2. **New products and services:** the ability to gauge information users' needs and satisfaction through analytics would improve the power to give library users what they want and satisfy their information needs. Davenport and Kim (2013) stressed that with Big Data analytics, more establishments including libraries are creating new products to meet users' needs.
3. **Faster, better decision making:** decisions can be made based on what they have learned. For instant, with the speed of Hadoop and in-memory analytics, combined with the ability to analyze new sources of data, libraries can analyze information at a faster rate.

According to Panda (2021), Big Data analytics can be applied in the following areas of the libraries:

1. **For superior search results:** data mining and text analytics on the past loan records and book bibliographies could enhance search results and recommendations.
2. **Demand Analysis:** it would help forecast demands for new and existing titles.
3. **Planning library collections:** the technology used would optimize planning the category mix in the collections by taking into consideration the space and budget constraints. Libraries play the role of research assistance and help researchers in improving their data analytics skills by assisting them to conceptualize how data is searched and teaching them how to use different search terms to derive the maximum efficient result.
4. **Data management plan:** according to Panda (2021), librarians' database design and development skills can prove useful for organization and data mining processes in Big Data. Assisting with data management plans is one of the most common services offered by academic libraries (Panda 2021, p. 51).

Big Data has enormous benefits for various organizations including libraries, especially academic libraries.

Benefits of Big Data in Academic Libraries

The raw scholarly data is useful for the extraction of valuable information and the data extraction has three demanding conditions: accuracy, insurance, and scalability. As the accuracy of data extraction techniques directly influences statistical quality, it is vital to obtain accuracy as much as possible. Coverage depends on precision, which is as important as extracting proper systems. Scalability is task-specific to big scholarly data due to the huge amount of facts to be processed (Khan et al., 2017). As part of the mandatory functions of the library, it would make sense for the library to find a tool to store the data, another one for indexing it, and yet another to run queries against it. Additionally, to those traditional features, other functions could be added. For instance, with library big data, librarians might be able to advise local business owners seeking market information, assist students to run statistics for a project, and help researchers manage huge datasets effectively. According to Kaladhar, Naick and Rao (2018), in the library two aspects of data mining could be achieved: one is using data stored in the library and another is using the data collected during the process when users use the library service.

Big Data is divided into three (3) different parts called structure, semi-structured, and unstructured data, as shown in Figure 2.

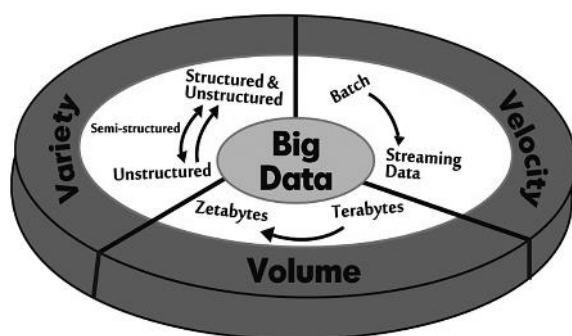


Figure 2: The 3 Major Parts of Big Data.

1. **Structured Data:** this is stored in the RDBMS data warehouse formally. Data are grouped in the form of rows and columns. About 10% of structured data is available around us. Structured data gives names to each field in a database and defines relationships between the fields, for example in library automation. Now most libraries are automated and bibliographic details of books and journals, with information about borrowers available in a structured form (Roy & Mukhopadhyay, 2023). Bibliographic information of a book entails the title, author, year, name of publisher, place of publication, ISBN, subtitle, edition, and number of pages. Other information like accession number, class, keywords, table of contents, prices, and details of supplier could be added. Also, for journals, there is the title, publisher, volume, issue, ISSN, Department, date published on, date received on, item code, supplier details, etc. Details of printed journals and information of borrowers include the borrower's identity, name, membership start date and end date, birth date, gender, email, contact number, address, blood group, etc. The collection of each library grows rapidly, thus structured data increases and it leads to Big Data. Libraries being information handlers and technology adopters (Garoufallou & Gaitanou, 2021). Big Data technologies will certainly affect their context by facilitating delivery of library services. In addition, Big Data protects each library user's right to privacy and confidentiality concerning information sought or received and resources consulted, borrowed, acquired, or transmitted and their bio-data information.
2. **Semi-structured Data:** this lies somewhere between structured and unstructured. It is not organized in a complex manner that makes sophisticated access and analysis impossible; however, it may have information associated with it such as metadata tagging which allows elements contained to be addressed.

A word document is usually considered to be unstructured data. However, one can add metadata tags in the form of keywords and other metadata that represent the document content. This makes it easier for that document to be found when the user searches for those terms and the data would now become semi-structured. Examples of this kind of data are RSS feeds, XML, and format data. Semi-structured data arise under a variety of forms for a wide range of applications such as institutional repositories, scientific databases, digital libraries, online documentation, and electronic resources. An institutional repository is an online archive of an institution's scholarly papers, deposited by their author. It may contain a variety of research outputs of organizations such as datasets, administrative documents, course notes, learning objects, and conference proceedings. Research and academic libraries are uploading documents to repositories.

3. **Unstructured Data:** this represents around 80% of data. It often includes text and multimedia content. Examples include library e-mail messages, word processing documents, audio-video lectures, photos, scanned images, presentations, blogs, webpages, social media, and many electronic documents. We should note that while these sorts of files may have an internal structure, they are still considered “unstructured” because the data they contain doesn’t fit neatly in a database. A different type of unstructured and scattered big data is available in libraries and this variety makes big data sets more challenging to organize and analyze.

Big Data benefits are enormous, and it is also useful in other sectors. This includes education, banking and security, communication and media entertainment, and healthcare providers, while public libraries has benefited from the application of Big Data, using different analysis techniques to analyze huge amounts of information.

Big Data Analyzing Techniques

The following are the six methods of analyzing Big Data. According to Rodríguez-Mazahua et al. (2016) and Manyika et al. (2011) Big Data can be analyzed using the following numerous techniques.

1. **Applied mathematics techniques:** Big Data analyzing technique can be used for the decision taken on the type of relationships that exist between variables. This might have occurred accidentally, and relationships might result from certain underlying causative relationships. However, these techniques are used to reduce the probability of type I errors (false-positive errors) and type II errors (false negative errors).
2. **Data mining technique:** this is a technique used to analyze huge information repositories and discovers implicit but potentially valuable information. Analysis of large amounts of information from Big Data leads to the emergence of new techniques such as data mining. It is also known as knowledge discovery in databases (Han, Pei & Kamber, 2011). The purpose of data mining is to expose the hidden relationships and unknown patterns and trends by mining huge amounts of data (Sumathi & Sivanandam, 2006). Three techniques are used to analyze data mining, namely classical statistics, artificial intelligence, and machine learning (Girija & Srivatsa, 2006). Data mining in academic libraries is called bibliominig. This concept is used to track patterns, behavioral changes, and trends in library system contacts (Siguenza-Guzman et al., 2015).

3. Machine learning: it serves the purpose of transferring the observational data in a way that can assist in the prediction of any hidden data. This technique was proposed by Bu et al. (2012). Nowadays, machine learning is generally used as a technique that drives the market and sales of online shops, keeps out spam emails, organizes advertising systems, and builds a content recommender system to suggest targeted users. It also supports scientists and researchers in different fields of science to interpret Big Data to acquire applicable knowledge, especially in certain fields such as high energy physics and biology.
4. A signal process: Big Data challenges offer several opportunities for signal process analysis, wherever knowledge-driven applied and mathematical learning systems can be visualized, simplified, and disseminated as analyzed results. Traditional and smart signal techniques for processing can be applied, such as primary part analysis, wordbook learning, and compressive sampling. However, the technique could lead to stress on the time-data adaptive process. Also, the primary role of the signal process in Big Data analysis is to simplify the character and scope of growing knowledge in various fields of science, especially multidisciplinary signal processing approaches (Slavakis et al., 2014).
5. Bibliomining tools: this is a developing trend, which helps identify behavioral patterns among library users and employees, and usage patterns of data resources throughout the library (Nicholson & Stanton, 2006). Bibliomining is also used to provide helpful and necessary data for library management needs, specializing in problem-solving, despite being greatly dependent on information technology. According to Prakash, Chand and Gohel (2004), Bibliomining regularly offers a comprehensive summary of the library workflow to monitor employee performance, identify areas of deficiency, and predict future user needs.
6. Multidimensional visualization: this technique has been widely used as a data abstracting tool, a data discovery tool in high-dimensional databases, and a data awareness tool in the cognitive process (Jain et al., 2012). This technique enables researchers to replicate the character of information efficiently according to its features, relationships, trends, and constellations (Ma, Shang, & Yuan, 2012).

Librarianship as a profession is saddled with the responsibility of connecting information users with relevant information resources. Big Data analytics is capable of uncovering new challenges of information utilization, consequently helping a librarian to fulfill their role as an embedded librarian, both in theoretical and practi-

cal terms. However, despite the huge benefits of Big Data in libraries, it does have some challenges.

Challenges of Big Data in Libraries

1. **Lacking of Data Analysts:** the essential issue is that data analysts need not only the skills of statistics and computer science but also skills of domain knowledge and collaboration ability. Hence, the challenges faced by librarians are the ability to manage the information of big data. It seems that short-course training might not be sufficient to handle this situation (Kaladhar et al., 2018).
2. **Adopting Big Data:** big data comes in different fields but a lot of companies are not ready for it. A study revealed that more than half of organizations could not handle big data due to a lack of personnel and platform. Research of library big data is even much slower than that in other disciplines. The key reason is that digital libraries tend to be self-contained organizational units and librarians try to stay back from new technology.
3. **Technical Challenges:** big data involves techniques such as capturing, storing, processing, and presenting data. There are different types of data applicable to the library environment, and this might come in various formats bearing significantly different statutes. Data in the library have various types and might be in different statutes. Some data might be waiting for digitalization and some data might be redundant and unusable. On the other hand, a huge set of data often contains some dirty or false data. Therefore, removing such data needs some work. As a result of the heterogeneous nature and formats of research data, assimilating them can be somewhat tedious. Many types of research data are considerably less useful when they are in their raw stage than after they have had filters or other processes performed on them. In the library, acquisition of library resources without further processing through classification and cataloguing will make resources less usable. Those required jobs need a budget to build tools and for the provision of other supports.
4. **Budget Issues:** even though people have to understand the great benefit of using Big Data analysis, IT investments such as analytics servers, high-performance computing servers are needed to optimize its usefulness. It appears that most library administrations have not yet placed Big Data on the table due to shrinking budgets. Research data managed by projects are given less attention because of the challenge of human resources. Moreover, a lot of research data that were produced ten years ago are still analogue, as digi-

tizing these resources is not a simple task; rather they need a lot of funds, time, and personnel resources.

5. Privacy: big data is mining the data and discovering knowledge. There should be a privacy issue that is eminent in mining and discovering knowledge. On the other hand, new risks of system intrusions might arise due to the accessibility of a huge amount of data. Data security issues have not been well considered for library big data concepts.

Conclusion

Big Data is a catchphrase. The capability of Big Data has captured the attention of the world of libraries. Data must be processed with advanced tools to reveal significant information. Librarians and information professionals can play a big role in the universe of Big Data, as they have the competencies, knowledge, and service mentality to assist all categories of information users irrespective of their discipline. With the aid of these powerful analytics that big data technologies offer, librarians can examine the data in new ways, thus adding value to different services and programs in the libraries and information centers. Big Data is the new frontier, and it represents a remarkable prospect for those who know how to utilize it. With meticulous organizational skills and a distinctive ability to scrutinize through huge amounts of information, librarians will soon be unique assets in the analysis and use of big data. The adoption and use of Big Data and advanced analytics are valuable and allow efficiencies in terms of cost, output, and innovation. Big Data has come to stay and librarians should continue to embrace the innovation for the librarianship profession to flourish.

References

- Anderson, J., & Rainie, L. (2018). The future of well-being in a tech-saturated world. Alvarado, R. U. (2016). Growth of literature on Bradford's law. *InvestigaciónBibliotecológica: Archivonomía, Bibliotecología e Información*, 30(68): 51-72.
- Alvarado, R. U. (2016). Growth of literature on Bradford's law. *InvestigaciónBibliotecológica: Archivonomía, Bibliotecología e Información*, 30(68): 51-72.
- Azeroual, O., & Fabre, R. (2021). Processing big data with apache hadoop in the current challenging era of COVID-19. *Big data and cognitive computing*, 5(1): 12.
- Begoli, E., & Horey, J. (2012). Design principles for effective knowledge discovery from big data. In *2012 Joint Working IEEE/IFIP Conference on Software Architecture and European Conference on Software Architecture* (pp. 215-218). IEEE.

- Bhoola, K., Kruger, K., Peick, J., Pio, P., & Tshabalala, N. A. (2014). Big data analytics. In *Proceedings of the actuarial society of South Africa's 2014 Convention* (pp. 106-144).
- Blascheck, T., Burch, M., Raschke, M., & Weiskopf, D. (2015). Challenges and perspectives in big eye-movement data visual analytics. In *2015 Big Data Visual Analytics (BDVA)* (pp. 1-8). IEEE.
- Borkar, V., Carey, M. J., & Li, C. (2012). Inside Big Data management ogres, onions, or parfaits? In *Proceedings of the 15th international conference on extending database technology* (pp. 3-14).
- Cai, L., & Zhu, Y. (2015). The challenges of data quality and data quality assessment in the big data era. *Data Science Journal*, 14, 2-2.
- Cartledge, C. (2016). How many Vs are there in big data <https://www.scirp.org/reference/referencespapers?referenceid=2853019>
- CISCO. (2012). Cisco's VNI forecast projects the internet will be four times as large in four years. <https://newsroom.cisco.com/press-releasecontent?type=webcontentandarticleId=888280>.
- Cukier, K. (2010). Data, data everywhere. *Economist*, 394 (8671): 3-5.
- Darvazeh, S. S., Vanani, I. R., & Musolu, F. M. (2020). Big data analytics and its applications in supply chain management. *New trends in the use of artificial intelligence for the industry 4.0*, 175.
- Data, IBM What Is Big. (2012). Bring Big Data to the Enterprise.
- Dave, P. (2013, October 21). Big data-data mining with hive-what is hive-what is hiveQL. <http://bolog.sqlauthority.com/2013/10/21/big-data-data-mining-with-hive-what-is-hive-what-is-hiveql-hql-day-15-of-21/>
- De Mauro, A., Greco, M., & Grimaldi, M. (2016). A formal definition of Big Data based on its essential features. *Library Review*, 122-135.
- Ferkoun, M. (2014). Cloud computing and big data: An ideal combination. <https://www.ibm.com/blogs/cloudcomputing/2014/02/cloud-computing-and-big-data-anideal-combination>.
- Garoufallou, E., & Gaitanou, P. (2021). Big data: opportunities and challenges in libraries, a systematic literature review. *College and Research Libraries*, 82(3), 410.
- Girija, N., & Srivatsa, S.K. (2006). "A research study: Using data mining in knowledge base business strategies. *Information Technology Journal*, 5(3), 590-600.
- Gordon-Murnane, L. (2012). Big Data: A big opportunity for librarians. *Online (Weston, CT)*, 36(5), 30-34.
- Han, J., Pei, J., & Kamber, M. (2011). *Data mining: concepts and techniques*. Elsevier.
- Herodotou, H., Lim, H., Luo, G., Borisov, N., Dong, L., Bilgen Cetin, F., & Babu, S. (2011). Starfish: A self-tuning system for Big Data analytics. *Cidr*, 11, 261-272.
- Jain, A. K., Narasimha M., & Flynn, P. J. (2012). Data clustering: A review. *ACM computing surveys (CSUR)*, 31(3), 264-323.
- Jewell, D., Barros, R. D., Diederichs, S., Duijvestijn, L. M., Hammersley, M., Hazra, A., Holban, C., et al. (2014). Performance and capacity implications for big data. IBM Redbooks, P.48, ISBN-13: 9780738453583
- Jiang, W., & Luo, J. (2022). Big data for traffic estimation and prediction: A survey of data and tools. *Applied System Innovation*, 5(1), 23.

Kaladhar, A., Naick, B. D., & Rao, K. S. (2018). Application of big data technology to library data: A review. *International Journal of Library and Information Studies*, 8(2), 25-30.

Khan, S., Liu, X., Shakil, K. A., & Alam, M. (2017). A survey on scholarly data: From big data perspective. *Information Processing and Management*, 53(4), 923-944.

Koga, S. (2015). Major challenges and solutions for utilizing big data in the maritime industry". https://commons.wmu.se/cgi/viewcontent.cgi?params=/context/all_dissertations/article/1489/&path_info=MSEA_2015_Koga.pdf

Laney, D. (2001). 3D data management: Controlling data volume, velocity and variety. *META Group Research Note*, 6(70), 1.

- Leonelli, S., & Tempini, N. (2021). Where health and environment meet: the use of invariant parameters in big data analysis. *Synthese*, 198(10), 2485-2504.
- Lomotey, R. K., & Deters, R. (2014). Towards knowledge discovery in big data. In *2014 IEEE 8th International Symposium on Service Oriented System Engineering* (pp.181-191).
- Ma, C.-L., Shang, X.-F., & Yuan, Y.-B. (2012). A three-dimensional display for big data sets. In *2012 International Conference on Machine Learning and Cybernetics*, vol. 4 (pp. 1541-1545). IEEE.
- Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C., & Hung Byers, A. (2011). Big data: The next frontier for innovation, competition, and productivity. McKinsey Global Institute. p.143, ISBN 0983179697, 9780983179696
- McGuire, T., Manyika, J., & Chui, M. (2012). Why big data is the new competitive advantage. *Ivey Business Journal*, 76(4), 1-4.
- Morris, R. J. T., & Truskowski, B. J. (2003). The evolution of storage systems". *IBM Systems Journal*, 42(2), 205-217.
- Normandeau K (2013). Beyond Volume, Variety and Velocity is the Issue of Big Data Veracity. <https://insidebigdata.com/2013/09/12/beyond-volume>
- Oguntimilehin, A., & Ademola, E.-O. (2014). A review of big data management, benefits and challenges. *A Review of Big Data Management, Benefits and Challenges*, 5(6), 1-7.
- Olabode, O. E., Boso, N., Hultman, M., & Leonidou. C. N. (2022). Big data analytics capability and market performance: The roles of disruptive business models and competitive intensity. *Journal of Business Research*, 139, 1218-1230.
- Olendorf, R., & Wang, Y. (2017). Big data in libraries. In S. C. Suh & T. Anthony. (Eds.), *Big Data and Visual Analytics* (pp. 191-202). Springer.
- Oussous, A., Benjelloun, F.-Z., AitLahcen, A., & Belfkih, S. (2018). Big Data technologies: A survey". *Journal of King Saud University-Computer and Information Sciences*, 30(4), 431-448.
- Pan, W., Li, Z., Zhang, Y., & Weng, C. (2018). The new hardware development trend and the challenges in data management and analysis. *Data Science and Engineering*, 3(3), 263-276.
- Panda, S. (2021). Usefulness and impact of big data in libraries: An opportunity to implement embedded librarianship. *Technological Innovations and Environmental Changes in Modern Libraries*, 45-60.
- Prakesh, K., Chand, P. & Gohel, U. (2004). Applications of Data Mining in Library and Information Services, 2nd Convention PLANNER - 2004, Manipur University, Imphal 4-5 November, pp 169-177
- Power, D. J. (2014). Using 'big data' for analytics and decision support". *Journal of Decision Systems*, 23(2), 222-228.
- Ristevski, B., & Chen, M. (2018). Big data analytics in medicine and healthcare. *Journal of Integrative Bioinformatics*, 15(3).
- Rodríguez-Mazahua, L., Rodríguez-Enríquez, C.-A., Sánchez-Cervantes, J. L., García-Alcaraz, C., & Alor-Hernández, G. (2016). A general perspective of Big Data: applications, tools, challenges and trends. *The Journal of Supercomputing*, 72(8), 3073-3113.

Roy, B. K., & Mukhopadhyay, P. (2023). Theoretical backbone of library and information science: A quest. *LIBER Quarterly: The Journal of the Association of European Research Libraries*, 33(1), 1-57.

Scott, D. W. (2015). *Multivariate density estimation: Theory, practice, and visualization*. 2nd edition, P.384, John Wiley & Sons, ISBN: 9781118575536

Sengupta, S. (2016). Big Data: The next big opportunity for librarians, 89-91.

Siguenza-Guzman, L., Saquicela, V., Avila-Ordóñez, E., Vandewalle, J., & Cattrysse, D. (2015). Literature review of data mining applications in academic libraries. *The Journal of Academic Librarianship*, 41(4), 499-510.

- Sivarajah, U., Kamal, M. M., Irani, Z., & Weerakkody, V. (2017). Critical analysis of big data challenges and analytical methods. *Journal of Business Research*, 70, 263-286.
- Slavakis, K., Kim, S. J., Mateos, G., & Giannakis, G. B. (2014). Stochastic approximation vis-a-vis online learning for big data analytics [lecture notes]. *IEEE Signal Processing Magazine*, 31(6), 124-129.
- Sonawane, A. S. (2018). Big data: Application in libraries. *International Journal of Scientific Research in Multidisciplinary Studies*, 4(1), 22-23.
- Sumathi, S., & Sivanandam, S. N. (2006). Text and Web Mining. *Introduction to Data Mining and its Applications*, 545-589.
- Syed, A. (2015). Data science is emerging as a hot new profession and academic discipline.
<http://www.personstyle.com/category/data-science-practices/>
- Teets, M., & Goldner, M. (2013). Libraries' role in curating and exposing big data. *Future Internet*, 5(3), 429-438.
- Ward, J. S., & Barker, A. A. (2013). Undefined by data: A survey of big data definitions. [https://doi.org/ 10.48550/arXiv.1309.5821](https://doi.org/10.48550/arXiv.1309.5821)
- Witt, S. W., Dimitrios, A., Koutsomitropoulos, G. D. M. Y. A., Solomou, R. J., Spires, M., Cohen Miller, A. S. et al. (2018). Co-authorship and diversity among researchers: Positive trends for IFLA and the global profession. *Co-authorship and diversity among researchers: Positive trends for IFLA and the global profession*, 3.