

Big Data Applications the Banking Sector: A Bibliometric Analysis Approach

SAGE Open
October-December 2021: 1–18
© The Author(s) 2021
DOI: 10.1177/21582440211067234
journals.sagepub.com/home/sgo

Haitham Nobanee^{1,2,3} , Mehroz Nida Dilshad^{3,4},
Mona Al Dhanhani¹, Maitha Al Neyadi¹, Sultan Al Qubaisi¹,
and Saeed Al Shamsi¹

Abstract

This study aims to review the existing literature on big data applications in banking using a bibliometric analysis approach. This approach describes citation rates, research outputs, and their implementations, along with current streams in the field and future research agenda. The articles were selected from 2012 to 2020 and sorted by the citation rate in results and analysis. We have discovered 60 papers related to big data in banking, although the applications of big data in the banking sector are growing rapidly, the number of research output in this field is limited. Several themes are extracted from the studies that are reviewed, analyzed, and presented in this report. This review covered the themes that include investment, profit, competition, credit risk analysis, banking crime, and fintech. This report also signifies the importance, use of big data, and its function in the banking and financial sector. This study has also discussed the future research scope in the banking industry's big data analytics.

Keywords

Big Data, finance, banking, anti-financial crime, bibliometric

Introduction

International organizations contribute immensely in the global arena. According to Abbott and Snidal (1998), international organizations help manage several global concerns involving monetary and health policies worldwide. It is an institutional agreement among members or participants of a worldwide system to ensure the achievement of goals that were established and aligned with aspirations, reflecting attributes, systematic conditions, and concerns of the international organization's members (Hanrieder, 1966). The sovereignty of the nation-state provides the fundamental rule of them (Barkin & Cronin, 1994). The term big data refers to massive data sets containing information on different fields of business, human behaviors, industries, and many more. Big data uses specific techniques and methods to analyze the extensive data set to conclude meaningful results or to predict future trends and directions. Sagioglu and Sinanc (2013) explained that big data is challenging to store, analyze, and visualize for future investigation and outcomes. Various structures, complex correlations, and hidden patterns in big data make information gathering difficult for organizations. However, the audience can gain deeper insights from these enormous amounts of data to gain a competitive advantage over companies in the market. Interestingly, cookie data and social media analytics also significantly gather massive

amounts of big data for firms. Marx (2013) highlighted an issue in their study, stating that sharing big data remains a challenge for customers and collaborators.

Big data comprises unstructured, semi-structured, and structured data mined by organizations to obtain helpful information to improve operations and increase profitability levels. It is implied that big data is gaining recognition widely by decision-makers for strategy-driven results (Labrinidis & Jagadish, 2012). It is also suggested that big data can broadly transform businesses' processes (Mishra et al., 2018). Big data sets are usually complex as compared to other data sets. Still, it provides more information and insight regarding the specific field and guides the management toward better future decisions (Marr, 2016). Big data holds significant importance in financial institutions and banks as it helps to make effective decisions and develop effective policies. Structured and unstructured data is used to

¹Abu Dhabi University, UAE

²University of Oxford, UK

³The University of Liverpool, Merseyside, UK

⁴Curtin University, Dubai, UAE

Corresponding Author:

Haitham Nobanee, College of Business, Abu Dhabi University, Al Ain Road, Abu Dhabi 59911, UAE.

Email: nobanee@gmail.com



Creative Commons CC BY: This article is distributed under the terms of the Creative Commons Attribution 4.0 License (<https://creativecommons.org/licenses/by/4.0/>) which permits any use, reproduction and distribution of

the work without further permission provided the original work is attributed as specified on the SAGE and Open Access pages (<https://us.sagepub.com/en-us/nam/open-access-at-sage>).

create better strategies and to anticipate customer behavior (Sun et al., 2014). We have used a bibliometric analysis to provide an in-depth understanding of the international institution's role in the global arena (Winther et al., 2009). The procedure of bibliometric ensures how theoretical inquiry relationships can be formed Reynolds et al. (2006). It is essential to note that bibliometric review requires several techniques and software. Initially, dimensions databases were used in this study to gather relevant articles and journals, followed by a VOS viewer where dimensions database files were exported to create citation mapping. Finally, the contextual analysis was presented through tables and figures.

Literature Review

Business intelligence's evolution began with the extraction of valuable information from historical data. The extraction process was achieved by the construction of data storage archives to effectively manage and analyze the available data with speed and in vast layouts (Dicuonzo et al., 2019). Business management has been complexed with advancements in internationalization and innovation because business administrators require better support regarding their decision-making policy (Srivastava & Gopalkrishnan, 2015). Hence, data-driven decisions' importance and value are increasing gradually (Sun et al., 2014). When it comes to big data analytics trends, there are two significant themes in companies. The first theme includes companies relying on big data analytics to analyze or recognize new opportunities. Furthermore, it is also used to improve current products or services and optimize internal processes (Hung et al., 2020).

On the other hand, new companies mostly rely on big data analytics to create innovative products and services. The banking industry also acts as an early stage regarding data-driven decision-making (Hassani et al., 2018a). When it comes to big data, it can be described as "large volumes of high velocity, complex and variable data that require advanced techniques and technologies to enable the capture, storage, distribution, management, and analysis of the information" (Hassani et al., 2018a). Such a tool has given firms, worldwide, the opportunity to investigate the large amounts of data they hold for anomalies or unfamiliar information that are beneficial in decision-making practice. Previous studies concluded five traits of big data: volume as in the data quantity, rapidity of the data generation, value, accuracy, and variability of data formats including structured, semi-structured, and unstructured (Dicuonzo et al., 2019).

Big data analytics has proven to be a critical strategy to be implemented within the banking industry. Such significance is due to the growth of the collected and stored operational data, as it spikes the value of big data in the analysis of various internal business processes (Rahman et al., 2019). It is to be noted that huge, stored data requires a more sophisticated big data mining tools installed. The standard big data software can only analyze 5% of the entire bank's dataset at

a time (Rahman & Iverson, 2015). Therefore, it is crucial for, in this case, banks to invest in an advanced data mining software to achieve an accurate analysis of bank's data and potential risks. For example, Hadoop is a renowned big data software common among service providers to analyze their historical data (Bhuvana et al., 2016).

An important function of big data analytics (BDA) in the banking sector is its support in the development and visualization of various indicators of multiple internal and external processes' ineffectiveness (Bhuvana et al., 2016). Such indicators recognition, for one, is in the case of fraudulent activities. For example, the implantation of big data within banks could aid in the detection of fraud within occurring monetary transactions and in the prevention of negative consequences (Bhuvana et al., 2016). Additionally, big data allow banks to cater to their clients various customized products that inclines their performance rate (Bhuvana et al., 2016). Furthermore, Al-Dmour et al. (2021) concluded in their study that the application of BDA will positively impact banks' performances. Conclusively, Previous research group the big data functions in the banking industry under three categories: "customer relationship management (CRM), fraud detection and prevention, and risk management and investment banking" (Dicuonzo et al., 2019).

Banks hold and store a large amount of customer's transactional, behavioral, and demographics data. Therefore, the value and usage of big analytics have proved to be vital, especially regarding marketing improvement in commercial banks and risk management performance (Lee, 2017). However, the corporate banking sector acts as a primary source of revenue for the whole banking sector; therefore, applying analytics in corporate banks' marketing has not gained much popularity (Cook et al., 2012). On the other hand, corporate banking involves a higher transactional amount per customer than personal banking, such as personal loans versus corporate loans. Resultantly, the focus and attention of research efforts are primarily on risk management within the company.

Numerous research studies have concluded that corporate banking's central aspect is its relationship with customers. There are some reasons behind the new risk-aversion trend. It has been considered that the value and profit in corporate banking are high because businesses collaborate with various banks simultaneously, thereby resulting in opportunities to expand the business further and increase the profitability ratio. Moreover, customers in the corporate banking sector are highly valued compared to a client of commercial banks. Consequently, the focus and attention of research are on risk management so that corporate customer remains valued. Several mechanisms have been included regarding management and monitoring so that information asymmetry can be reduced. Also, high-risk corporations can be excluded because of increased monitoring. Big analytics and advanced analysis methods are gaining popularity along with the usage of non-traditional data. This non-traditional data involves a

recent adoption of non-financial data. The reason behind this is to improve the dynamic monitoring or credit evaluation further.

Advancement in technology has allowed some banks to shift data analytics to remain competitive and deal with the challenges of financial markets. An advantage of using data analytics for banks is having customers' historical data if needed. Conferring resource-based theory, understanding the business's collected or collectible data resources, and finding a proper way to use them is key to gaining a competitive advantage over other companies (Marr, 2016). Evidently, banks incorporating big data analytics are gaining a 4% competitive advantage over others within the banking sector (Dicuonzo et al., 2019).

Big data analytics is a powerful fintech tool introduced in the market to support business sustainability initiatives. Specifically, BDA allows businesses, including banks, to improve their operation internally. For example, banks could determine potential risk predictors and identify critical factors that could increase their customer retention rate, thus ensuring the sustainability of a business in the current excessively competitive market. Wang et al. (2021) highlighted the benefits of implementing fintech, including BDA, in commercial banks that encompass improved business models, reduced operating costs, enhanced services' efficiency, advanced risk control strategies, and increased competitiveness. However, when banks and other businesses decide to implement BDA in their operation, they must first lay down the foundation, that is, the equipment, that their success will rely on. This foundation consists of BDA hardware equipment and advanced big data software. The required hardware consists of "information network facilities, high-performance computers and cloud servers, and large-capacity storage" (Wang & Sui, 2020). Additionally, the software must display the following traits: great data mining capabilities, sophisticated artificial intelligence, spread storage, and set processing (Wang & Sui, 2020).

Method

Marín-Marín et al. (2019) reiterated that analyzing big data to reveal specific patterns have gained popularity in the last decade. Liu et al. (2019) further elaborated that extensive research on this area of interest presents immense potential for academia and businesses. Kalantari et al. (2017) demonstrate the bibliometric approach's value in their study by revealing the latest research trends of big data in different domains. Similarly, (Xian & Madhavan, 2014) used bibliometric analysis to unveil scholarly collaborations on big data topics amongst researchers. In response to the rapid growth of research on big data, this study has applied bibliometric analysis to compare the most significant and influential research contributions by researchers made toward this topic (Nobanee et al., 2021). This study will equip the audience to understand and harness the benefits from the information

collected through big data analysis. With the help of bibliometric analysis coupled with network analysis, a comprehensive literature review on big data has been presented for almost 40 years (1960–2020). This research offers insights on scholars' quality contributions and the latest advancement toward this topic in scientific journals. Furthermore, bibliometric analysis has helped identify critical countries, authors, and research clusters related to this study area (Nobanee, 2020).

Different methods have been used in this bibliographic review to analyze the data. In this study, articles between 1960 and 2020 are filtered by using the Scopus search engine. Most of the relevant articles have been published after 2005. Research papers have been sorted based on the citation rate, which depicts how other studies have referenced a study. The highest citation rate shows that the study has received the most significant attention from researchers because of its valued content. Data in this study has been demonstrated through different tables and figures based on the citation rate. In addition, this study shows a graphical representation of documents, author, co-author, citation, link strength, application organization, and country of extracted articles (Nobanee, 2021).

Keywords and Initial Search

Initially, "big data*" and "bank*" search keywords were used to cover the whole essence of this topic and collect data. These combinations have been made to ensure that significant keywords get the desired results in documents. Sixty documents were initially shown in the Scopus database. This database was chosen because it is the most critical source of peer-reviewed journals from Emerald, Taylor and Francis, Inderscience, Springer, and Elsevier publishers (Fahimnia et al., 2015). In addition, Scopus is a more comprehensive database for journals than the Web of Science (WoS) database (Yong-Hak, 2013).

Refined Search

To gather the most relevant papers, the authors searched the titles of the document using the advanced query that included the primary search keywords (big data and bank*) with the addition of the limitations and exclusions to the search, shown in Table 1. The search was limited to documents written in English, the analysis period between 2012 and 2020, the source type to conference proceedings, journal articles, and book series. The search results show 80 papers. After conducting a manual cleaning of irrelevant documents such as those related to filter banks, blood banks, tissue banks, riverbanks, fertility banks, etc. and the exclusion of earth and planetary sciences subject area, the results included 60 papers. The following refinement were included to eliminate the inclusion of irrelevant papers to this study, as it focuses on the application of big data in the banking sector solely and

Table 1. Refined Search Query.

Description	Conditions
Search query	(TITLE (big AND data) AND TITLE (bank*)) AND (LIMIT-TO (LANGUAGE, "English"))
Search query after refining	(TITLE (big AND data) AND TITLE (bank*) AND NOT TITLE (milk AND bank) AND NOT TITLE (food AND bank*) AND NOT TITLE (riverbank) AND NOT TITLE (data AND bank) AND NOT TITLE (tissues AND bank) AND NOT TITLE (fertility AND bank) AND NOT TITLE (blood AND bank*) AND NOT TITLE (bankura) AND NOT TITLE (filter AND bank) AND NOT TITLE (capacitor AND bank) AND NOT TITLE (battery AND bank)) AND (LIMIT-TO (PUBSTAGE, "final")) AND (LIMIT-TO (PUBYEAR, 2020) OR LIMIT-TO (PUBYEAR, 2019) OR LIMIT-TO (PUBYEAR, 2018) OR LIMIT-TO (PUBYEAR, 2017) OR LIMIT-TO (PUBYEAR, 2016) OR LIMIT-TO (PUBYEAR, 2015) OR LIMIT-TO (PUBYEAR, 2014) OR LIMIT-TO (PUBYEAR, 2012)) AND (LIMIT-TO (LANGUAGE, "English")) AND (LIMIT-TO (SRCTYPE, "j") OR LIMIT-TO (SRCTYPE, "p") OR LIMIT-TO (SRCTYPE, "k")) AND (EXCLUDE (SUBJAREA, "EART"))
Access	Both open access and non-open access documents are included.
Query search date	July 15th, 2021
Years	2012–2020
Subject area	All subject areas except earth and planetary sciences are included.
Source type	The search was limited to Conference Proceedings, Journals, and Book Series
Language	The search was limited to the English language.

none other. The period was restricted 2012 to 2020 because it's when a growth in publication on the topic occurs in the literature. Restricting the language to English is to universalize this paper for any English-speaking user, as the English language is today a universal language communicated amongst various countries and races. The sources were limited to three types that follow under the category of studies, which, for example, exclude books. Lastly, earth and planetary sciences subject area was excluded from the search query to prevent irrelevant documents on, for example, riverbanks appearing in the final collected data that decreases the usefulness, value, and accuracy of this paper.

Results

In this section, different metrics have been evaluated with the help of bibliometric tools to demonstrate and comprehend emerging research trends on this research topic.

Big Data Word Tree

Figure 1 represents a big data word tree covering several research areas of the banking system that different researchers have explored. This figure was constructed on NVivo software using the collected papers from Scopus. The trending banking-related aspects branched out in this tree chart are security, customer satisfaction, data, digital banking, fraud and crime, credit, marketing, services, bankruptcy, artificial intelligence. Researchers have discussed most of these topics in their research papers published in peer-reviewed journals.

Most Influential Authors

Table 2 lists down names of researchers who contributed high-impact research papers on big data in banking. The data

in the table is sorted according to the citation rate of research papers produced by authors. The most common topics explored by researchers include data-based banking, customer analytics on big data-based banking, big data-based individual bank card transaction, and big data of bank cards. These are the most frequently used keywords identified in these shortlisted publications. The highest citation rate of these articles ranges from 31 to 37, which denotes a significant citation rate and the worth of that topic in the research field.

The top three ranked authors have scored a 37-citation rate and published only two documents on the subject except for Morris contributing a single article to the topic's literature. Arias and Ratti worked on both papers that each published. The first discussed the importance of big data analytics on categorizing the mobility of Spanish bank's local and foreign customers in presenting new business opportunities (Sobolevsky et al., 2014). Their second publication identified a strong connection between "individual spending behavior and official socioeconomic indexes" that offer critical information on the quality of life of individual bank clients (Sobolevsky et al., 2017). The investigators recognized (Sobolevsky et al., 2014) paper had gained more attention than their 2017 article because the first accumulated a citation rate of 31 while six studies only cited the second. The reason for (Sobolevsky et al., 2014) paper's popularity is the achievement of great attention from papers studying the tourism sector, transportation systems, and customers' purchase behavior analysis.

On the other hand, Morris released a single article relating to the topic of interest that solely achieved 37 citations up to 2020. The paper constructed a model for analyzing massive, stored data to create value for banks from insights about clients and their behavior obtained from analysis and is known as the iCARE framework (Sun et al., 2014). The papers that

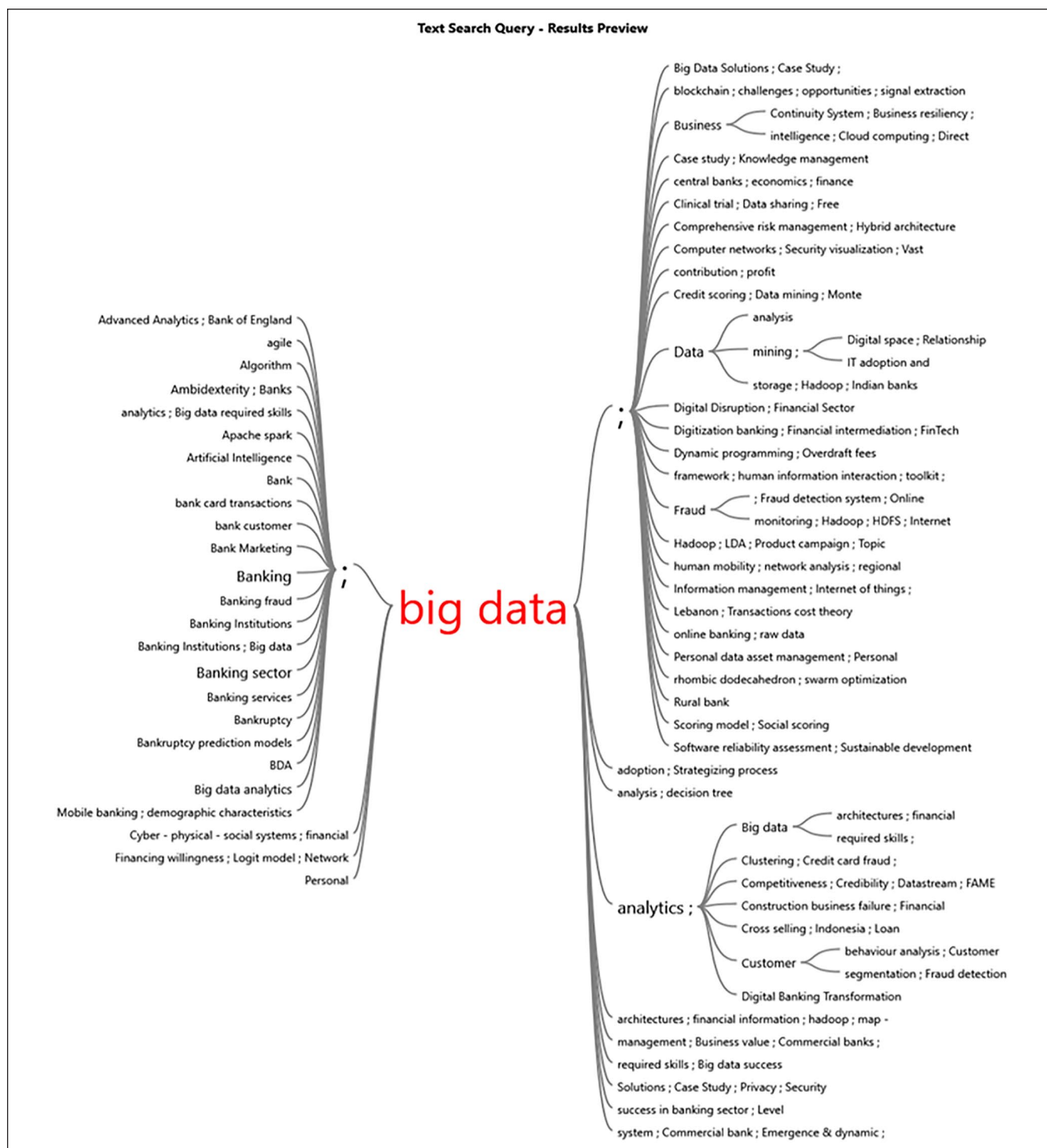


Figure 1. Big data word's tree (NVivo software).

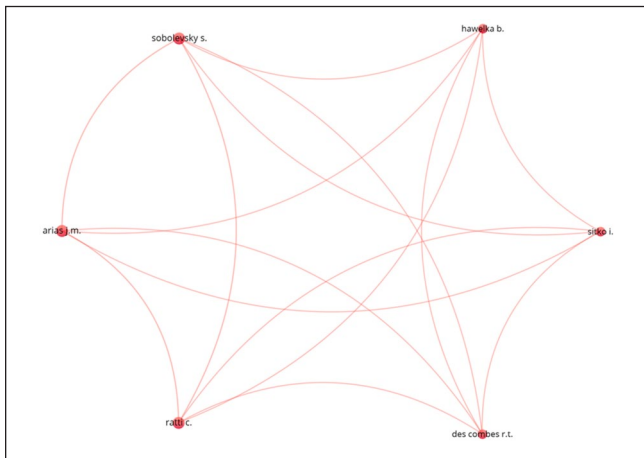
took an interest in Morris's study focus on analyzing vast aspects of the banking sector and the construction of effective business models.

Figure 2 presents an author's map used as an indicator to assess the research publication and its impact activity in this discipline. The clusters in this figure represent the author's

co-citation analysis of publications on big data in banking and relevant emerging themes. This figure depicts a collaboration between authors and is evident that the greatest collaboration with the most influence in the field focus on the (Sobolevsky et al., 2014) and the (Sobolevsky et al., 2017) papers.

Table 2. Most Influential Authors.

Rank	Author	Documents	Citations	Rank	Author	Documents	Citations
1	Arias J.M.	2	37	6	Xie M.	1	37
2	Morris J.G.	1	37	7	Xu J.	1	37
3	Ratti C.	2	37	8	Zhu X.	1	37
4	Sobolevsky S.	2	37	9	Des Combes R.T.	1	31
5	Sun N.	1	37	10	Hawelka B.	1	31

**Figure 2.** The bibliometric map of authors' network.

Most Productive Organizations

Table 3 presents information on the top 10 organizations co-authors have been associated with for their publication work. This data has also been sorted based on citations. The table shows that the highest citation rate is 37 of Senseable City Lab from the US, IBM global business services from the US, and IBM Research Center China. The other organizations established in different countries offering different specializations also have a higher citation rate. It is implied that they have recorded and published research on big bank data analytics in various domains.

The first affiliation released two documents that accumulated in total 37 citations. Apparently, these publications are those contributed by Arias & Ratti, discussed previously. The total citation is split among both articles: (Sobolevsky et al., 2014) was cited by 31 papers, while (Sobolevsky et al., 2017) was cited by six studies. IBM Global Business Services in the United States and IBM Research in China collaborated on the single research they released on big data in the banking sector, and it is (Sun et al., 2014) study that was cited by 37 articles and Morris worked on as mentioned in the previous section.

Figure 3 shows a cluster chart of 11 organizations of the co-authors which have contributed significantly to big data in banking research. This chart shows the name of organizations with the highest citation rate and link strength. The figure clearly illustrates the organizations that contributed

almost equally to the field with associations with fellow affiliations on the publications they issued. The partnerships presented are the most significant in the literature on big data and banking fields and are focused on a single paper between the 11 affiliations in the figure. This study discusses the aspect of big data in monitoring “photovoltaic (PV) solar electricity Systems” within Europe (Reinders et al., 2019). This study presents the vitality of big data applications in monitoring certain internal business activities and reducing risks and operational costs that could be adapted in the banking industry.

Co-author Networks and Collaboration

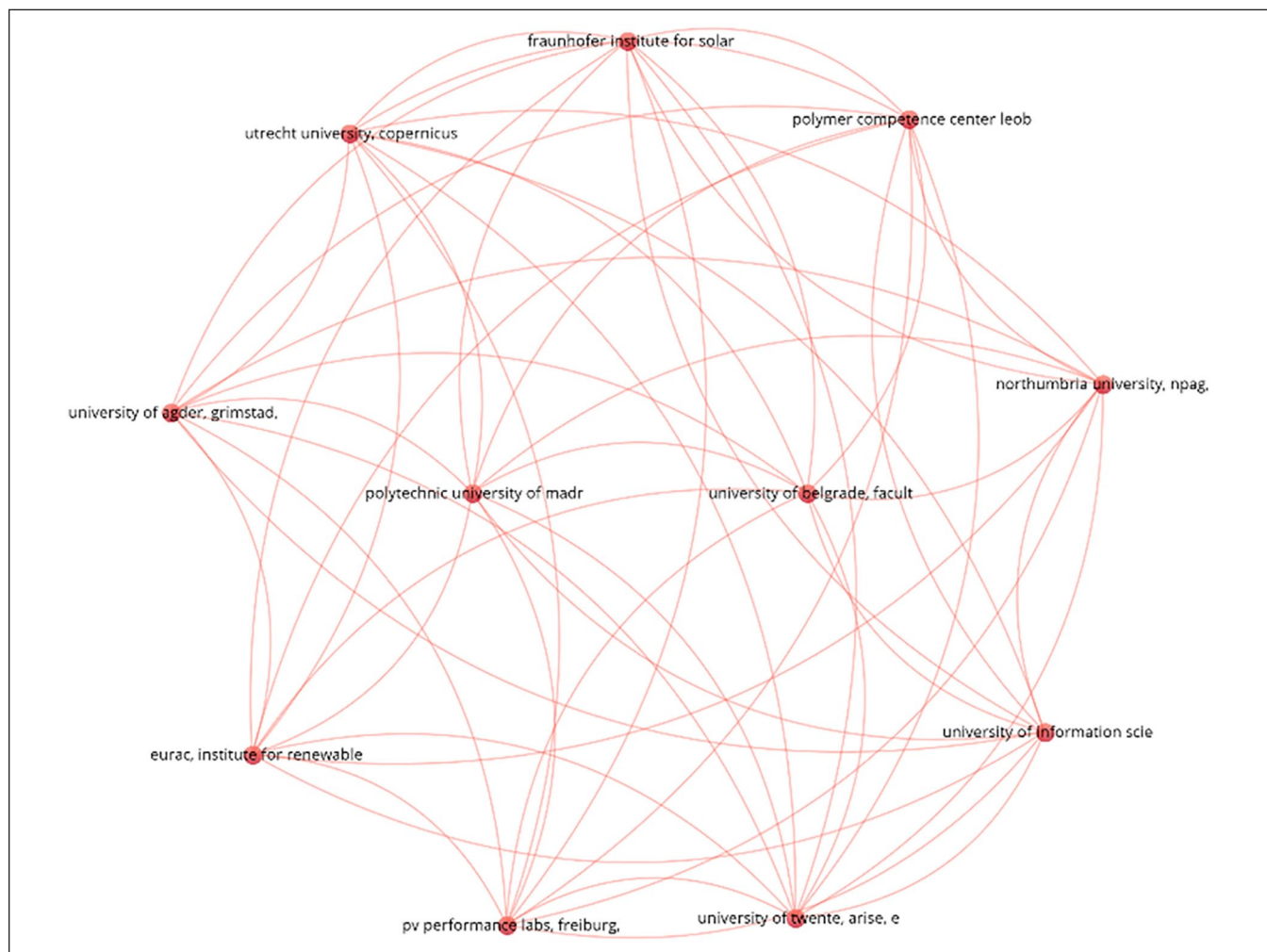
Co-authors to the author are those who also contribute to a research study and assist the principal author. This table has been sorted according to the highest citation rate of the coauthor's work. These applications include the big data of banking concerning customer analysis, credit card utilization, risk assessment, and competition. The highest citation rate with an adequate link strength is the publication of coauthor Arias J.M. whose publication was based on big data analysis of credit cards. Figure 4 represents different clusters of co-authors producing research work on various big data themes in banking, with a high citation rate. These clusters indicate a strong relationship between co-authors and authors for similar research interests. This figure also helps in comprehending an occurrence of research papers jointly produced by authors and co-authors. Nodes here indicate a relationship strength between researchers in this network.

Leading Countries

Table 4 illustrates the leading countries based on the number of publications in the field of interest: China, India, United States, United Kingdom, and Spain, who published 12, 11, 7, 6, and 5 papers, respectively, during the analyzed period. Following the analysis design of this paper, the authors will focus on exploring the three highest-ranked countries' studies. All 12 of china's-based articles focus on the application of big data in various scopes of its banking sector. China's pronounced interest is due to its rapidly emerging financial industry. This industry encompasses six different financial modes of operation, including a “big data-based financing mode” (Zhao et al., 2015).

Table 3. Most Productive Organizations.

Rank	Organization	Documents	Citations
1	Senseable City Lab, MIT, Cambridge, Ma, United States	2	37
2	IBM Global Business Services, New York, NY 10016, United States	1	37
3	IBM Research—China, Beijing, 100193, China	1	37
4	Department of Geoinformatics—Z-Gis, University of Salzburg, Salzburg, Austria	1	31
5	New Technologies, Bbva, Madrid, Spain	1	31
6	Symbiosis Institute of Business Management, Symbiosis International University, Lavale, Pune, 412115, India	1	23
7	Faculty of Business and Law, De Montfort University, Leicester, United Kingdom	1	16
8	London College of Fashion, University of The Arts London, London, United Kingdom	1	16
9	Research Institute for Energy Management and Planning, University of Tehran, Tehran, Iran	1	16
10	Department of Economic and Financial Studies, Miguel Hernández University of Elche, Spain	1	15

**Figure 3.** Network visualization of most influential organizations.

Most of India's research is devoted to determining the role of big data in banking, predicting bankruptcy using big data, using big data to identify profit stimulants, and analyzing their historical client-issued loans' data. The authors observed a collaboration between India and China on an

article that studies the "Impact of big data analytics on banking sector" (Srivastava & Gopalkrishnan, 2015).

The United States focused more on implementing fintech in the banking sector, including big data and artificial intelligence technologies, and studying the effective strategies of

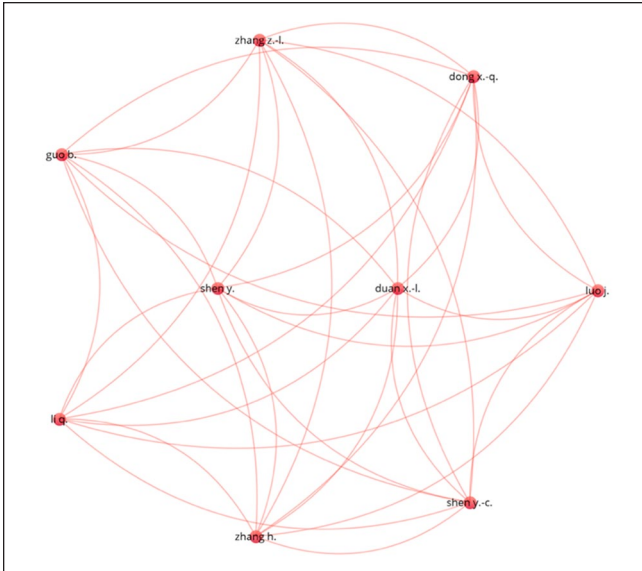


Figure 4. Co-author network and visualization.

Table 4. Leading Countries by Documents.

Rank	Country	Documents
1	China	12
2	India	11
3	United States	7
4	United Kingdom	6
5	Spain	5

Table 5. Leading Countries by Citations.

Rank	Country	Citations
1	United States	81
2	Spain	59
3	China	55
4	India	45
5	United Kingdom	45

applying big data, in specific, to the industry. The United States and China's collaboration is noteworthy, as they have constructed two big data and banking sector-related papers. The first focused on "Big data analytics for supply chain relationship in banking," and the second used "Big data analysis on demographic characteristics of Chinese mobile banking users" (Hung et al., 2020; Wang & Petrounias, 2017).

Instead, Table 5 presented the leading countries with contributions in the field of big data in banking. Based on the number of documents' citations, the United States led the list with 81 citations, followed by Spain with 59 citations, China with 55 citations, and India and the United Kingdom with 45 citations each. The high interest in United States' articles is because they mainly focus on various aspects and benefits of

implementing BDA. For example, they determine opportunities and challenges of big data implementation, fraud detection, customer behavior analysis, value creation, tourism, designing effective business models, economic performance prediction, and business innovation.

On the other hand, the 59 investigations are attentive to Spain's research. Its research encompasses big data topics including value creation, fraud detection, operation inefficiency examination, tourism, analysis of consumer behavior, designing business models, optimizing risk management efforts, predicting economic performance, and beneficial application in the financial field. Furthermore, China's-based documents interest researchers investigating several aspects and applications of big data, such as its application in finance, banking, industrial marketing, and credit card businesses. Additional interests fall down the line of enforcing social responsibility, presenting customer catering strategies, applying innovative technologies in the business field, increasing competitiveness, and inspecting the predictors of BDA adoption.

Figure 5 shows a network of countries authors have been affiliated made essential contributions to this research topic. This citation map further illustrates the count of the number of publications produced by countries, and these counts evaluate each country's contribution toward related research fields. The size of each node connected with other nodes represents the count of publications from each country. It is evident from the map that authors mainly associated China have made significant research contributions and collaborated with authors from other countries on this topic to elevate the big data in the banking sector, such as the United States, Norway, and Macao. China is followed by the United States, the United Kingdom, and Spain in influencing the big data in banking literature.

Most Cited Documents

Table 6 represents data of the research publications organized by citation rate between 2013 and 2019. Publications cover topics such as anti-fraud projects based on big data, monitoring banking systems with big data, bankruptcy prediction through the big data applications, or topology for human-centric banking big data. This table has filtered the top 10 publications from the gathered data and indicates their importance by the citation rate. Sun et al.'s (2014) paper created a model, known as iCARE, to analyze big customer data within banks. It is the most cited, accumulating 37 citations because it discusses a topic on managing customer data that will allow banks and businesses to benefit from such a strategy to increase their customer retention by adjusting their existing products or monitoring customer behavior. The second most cited article by 31 papers concentrates on analyzing clients' mobility through big data of transactions of bank's card (Sobolevsky et al., 2014). Its main aim is to investigate customer patterns and purchases through their

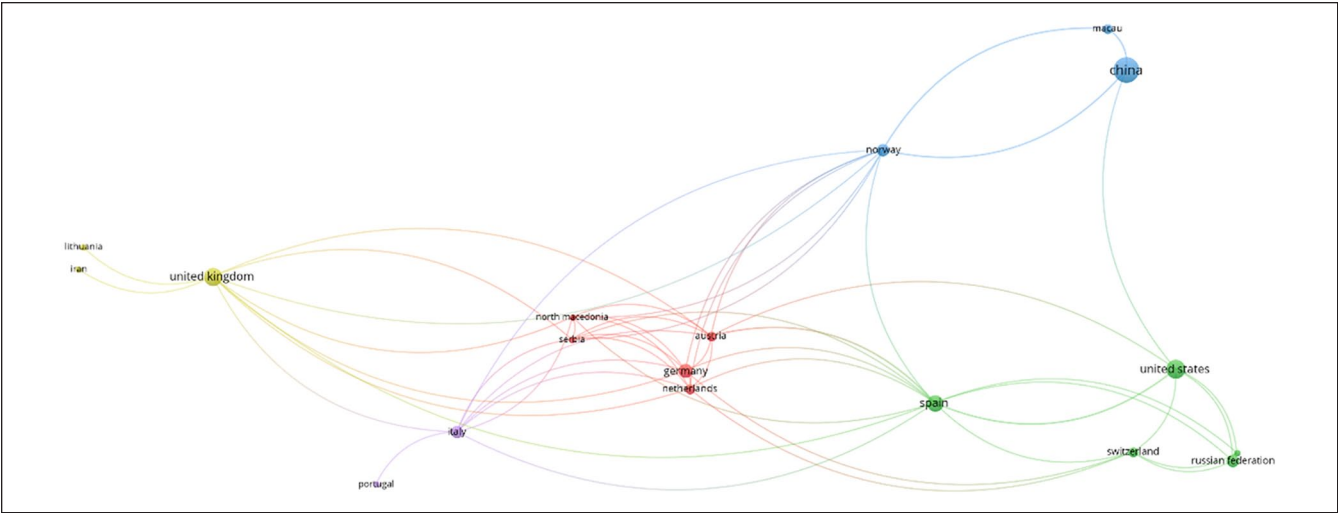


Figure 5. Leading countries.

Table 6. Most Cited Documents.

Rank	Document	Citations	Rank	Document	Citations
1	Sun et al. (2014)	37	6	Fuschi and Tvaronavičienė (2014)	13
2	Sobolevsky et al. (2014)	31	7	Fischer et al. (2015)	13
3	Srivastava and Gopalkrishnan (2015)	23	8	Evangelatos et al. (2016)	11
4	Hassani et al. (2018b)	16	9	Wang et al. (2019)	9
5	Pérez-Martín et al. (2018)	15	10	Hafiz et al. (2015)	9

historical transactions’ data to establish what location or store and product are popularly common amongst the bank’s clients. The banks could also benefit from such analysis to predict loan collection rates. Srivastava and Gopalkrishnan (2015) contributed the third paper, cited by 23 articles, generally conferring the influence of BDA on the Indian banking sector.

Figure 6 illustrates highly cited publications of five authors related to multiple aspects of bank big data from 2015 to 2018. This network shows the most influential documents produced by researchers across this subject area. As seen, the most intersected article to other contributed to the literature is Srivastava and Gopalkrishnan’s (2015) article mentioned earlier. It’s because the research encompasses the entire Indian banking sector, discussing vital aspects of big data’s effective integration, which could be applied with minor modifications in other banking sectors.

Keyword Frequency and Co-occurrence Analysis

This subsection discusses the most frequently used keywords in the relevant publications, shown in Table 7. For example, the most frequent keyword is the “big data” with 50 times occurrences. However, this table includes keywords with fewer occurrences, such as “big data analytics,” “sales,” “finance,” “artificial intelligence,” and “data handling.” Furthermore, the data in this table helps comprehend that big

data is an essential variable in the banking system as different publications have highly explored it. The reason is banks and financial institutions aim at managing their massive historical operational data effectively using fintech, such as big data and artificial intelligence technologies. These technologies provide the banking sector with the necessary tools to apply data mining and create value resulting from non-previously known success contributing elements.

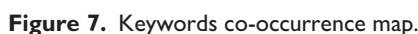
A cluster map in Figure 7 shows a broad diversification of keywords explored by different applications. The analysis results show that keywords such as “banking institutions,” “data technologies,” “information management,” “data mining,” “advanced analytics,” “risk management,” “the industrial revolution, are used by respective authors. It further shows that a wide range of topics under the big data in banking has been thoroughly explored to attain higher and improved performances of the banking sector. From the figure, the authors can conclude that most of the available documents revolve heavily around big data, as it explores the visualization aspect of banks” data to manage potential risks effectively, achieve a competitive market advantage, and ultimately sustain their profitability.

Most Relevant Sources

Table 8 presents a list of sources these documents have been traced from, that is, scholarly sources such as journals,



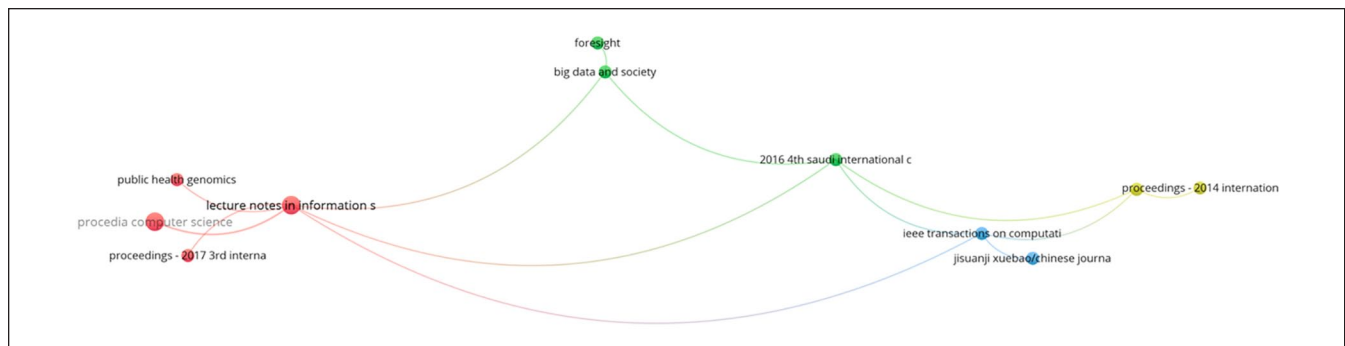
Rank	Keyword	Occurrences	Rank	Keyword	Occurrences
1	Big Data	50	6	Data Analytics	7
2	Big Data Analytics	9	7	Finance	7
3	Data Mining	9	8	Banking	6
4	Information Management	9	9	Artificial Intelligence	5
5	Sales	8	10	Data Handling	5



articles published on the topic are published by IBM Journal of Research and Development, Proceedings—2014 IEEE International Congress on Big Data, Bigdata Congress, and Procedia Computer Science and were cited 37, 31, 23 times, respectively. IBM Journal of Research and Development released Sun et al. (2014) articles that configured the iCARE big data model for banks. The 2 to 14 IEEE international big data congress published Sobolevsky et al.'s (2014) study on analyzing mobility of bank customers' transactions using BDA. Additionally, Procedia Computer Science released two

Table 8. Sources.

Rank	Source	Citations
1	IBM Journal of Research and Development	37
2	Proceedings—2014 IEEE International Congress on Big Data, Bigdata Congress	31
3	Procedia Computer Science	23
4	Journal of Management Analytics	16
5	Journal of Business Research	15
6	Information Visualization	13
7	Journal of Security And Sustainability Issues	13
8	Public Health Genomics	11

**Figure 8.** Sources network.

papers in the study's scope. The first disclosed big data to create solutions for security concerns in banking initiations (Salleh & Janczewski, 2019). Secondly, the article mentioned earlier on the big data adoption in the Indian banking sector by Srivastava and Gopalkrishnan (2015) was sourced by Procedia Computer Science.

Figure 8 represents a network of sources that have facilitated the publication of articles on big data in banking. These sources have a good impact factor. A significant number of sources based on various academic disciplines confirm their interest in publishing articles related to the banking sector. It is evident from the figure that the most influential article is sourced by Lecture Notes of Information System and Organization, which are two documents. This high recognition is due to multiple papers from several sources that have cited the two research contributed by Lecture Notes of Information System and Organization. The first article analyzes the implications of adopting big data in the Lebanese banking sector using the transactional cost theory (Chedrawi et al., 2020). In addition, the source released a second article that Unveils the big data adoption in banks by examining the strategies of implementing the new technology (Diniz et al., 2018).

Growth of the Publications

Figure 9 presents scholarly work produced by researchers on big data in banking from 2012 to 2019. It is noteworthy that the authors had the highest number of publications on this

topic in 2019. Thus, it depicts an increased awareness and significance of big data in the banking sector globally in recent years. On the other hand, the least number of research papers was published by researchers in 2012 and 2013. Furthermore, it shows that an effective and digitalized banking system has become a substantial need of the current time. Therefore, researchers need to delve deeper to analyze this topic in the banking industry. The applications of big data in banking will reshape the banking operation in the future as the banking sector is moving toward digitalization. Despite this topic's importance, the number of research output in this field is limited; specifically, the number of published research on the said topic is 60 documents in total. Hence, more analysis in big data in bunking should be carried on. The most significant publication rate was released in 2019, where 18 documents on the topic were contributed to the literature. Three out of the 18 papers received greater recognition, which is that of (Wanget al., 2019), (Hale & Lopez, 2019), and (Guha, 2019). Using the “fuzzy C-means algorithmGuha (2019) utilized BDA in predicting bankruptcy; thus, such a study becomes crucial to businesses, even banks, to prevent insolvency. Furthermore, Hale and Lopez (2019) studied the connectivity network between individual banks through big data for the possibility of integrating all banks” systems and become one whole sector network. Lastly, Wang et al. (2019) introduced a rhombic dodecahedron topology to efficiently gather data about bank's clients, which is insightful and valuable to currently operating banks worldwide.

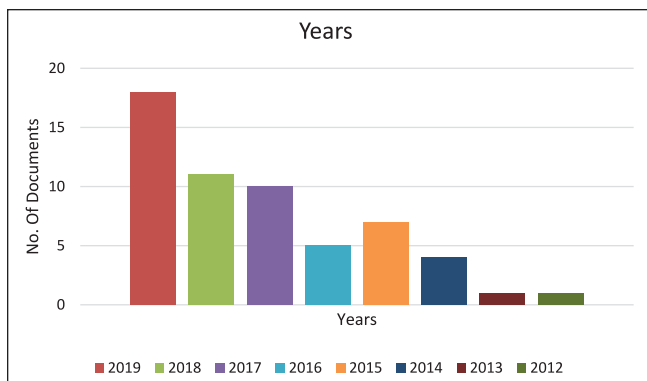


Figure 9. Number of documents published per year.

Research Themes

Most relevant articles have been filtered in Scopus, and their findings and future expectations have been listed in Table 9. Most importantly, these articles have been presented in different big data themes in the banking sector, that is, competition, profit, crime, supply chain, fintech, investment, fraud, and digital banking. A substantial amount of research has been carried out to bring more efficacy in the banking sector globally with the help of big data. Most of these articles prove that big data analytics plays a crucial role in multiple aspects of banking. These themes can be further explored in the future to fill more study gaps in this discipline. These articles summarized in the table below reveal emerging research areas of study within this topic. The following themes indicate that the magnitude of big data is growing in different sectors and enterprises. This massive volume of data is being analyzed in industries for various reasons to unveil facts to improve operations. Enterprises use this accumulated big data to gain valuable insight to compete with other companies, refine their techniques, make more informed decisions, and increase profitability. It is implied that big data analytics help enterprises gain a competitive advantage and become increasingly customer-focused.

Profit is another theme in which big data research papers have been filtered and shown in the content analysis table. Competition and increased profit levels demand unique, accurate, and massive processed data for business changes. Extracting such valuable information would enable organizations to explore new advantages and opportunities, leading to increased profitability. This table further points out another theme, crime, studied by researchers. It infers those big data analytics help to unfold crime patterns and facilitate investigations in banks. Researchers indicate that processed big data helps save immense costs, effort, and time in crime inquiry in the banking sector. Researchers in this area of research explore the supply chain theme, suggesting that digitalization is trending in the supply chain industry to improve operations significantly. Researchers have identified that finance firms can track evolving international

trends, improve strategies and increase efficiency levels. It is evident from research on this topic that the financial services sector has been revolutionized because of an increased value generated by big data for fintech, that is, financial technology companies. Credit risk assessment and scoring for banks have become affordable and quicker for fintech companies. It summarizes the essence of the fintech theme for big data research papers. Fintech companies in the banking sector now invest heavily in processed data to improve and innovate their services, compete with their competitors and increase customer confidence. The investment theme in this content analysis table covers the concept of investment driven by a large set of company or market data such as stock returns, prices, or publicly accessible financial statements.

Research suggests that this unconventional data help enterprises make more informed investment choices. Big data further helps banks and fintech companies obtain an informational advantage to uncover future opportunities. Large sets of data coupled with economic and statistical data are somewhat equipped with essential information for organizations. Investment companies root for quantitative data to process it efficiently and extract meaningful information. It is another positive theme on big data that has been trending in the research field and studied by researchers. Analyzing this structured data further enables companies such as banks to detect fraudulent activities. Tang and Karim (2019) reveal that dig data analytics help strengthen results, identify fraud indicators, and document current issues in banks. This fraud detection theme has also been recognized by researchers in the context of big data and listed in the content analysis table. It infers that big data has become a substantial part of fraud recognition algorithms in the current time. Another critical theme concerning big data is digital banking identified in this study. Banks are introducing new ways to advertise their products and services to facilitate customers' improved financial decisions. Big data composes a crucial role in customer retention and investing in new features to attract more customers for banks. Recent trends show that big data in digital banking is gaining importance to study customer's preferences, improve security services, and flag unusual transactions.

Discussion

After the conducted investigation through the 10 research themes, the researchers can conclude that big data plays a serious role in distinctive aspects of the banking sector. These scopes include the banks' competitiveness, profitability, internal and external criminal or fraudulent activities, implementation of fintech, and, but are not limited to, product improvement. As seen, BDA can form competitive business models for the benefit of the banking sector, such as a mathematical model that increase their market competitiveness. Additionally, BDA allows banks to explore consumers' responses on tried products, develop their existing products,

Table 9. Research Themes.

No.	References	Research question	Future expectations
Competition			
1	Hartmann and Nair (2010)	How is the dynamic structural demand model applicable in retail format? How retail format substitution model fit in an oligopoly retail environment in the banking system?	This study will guide the establishment of the retail bank format in a competitive environment.
2	Zhao et al. (2015)	How are evolutionary mathematical models of competition associated with big data-based e-commerce financial institutions and commercial banks? How are stable evolutionary strategies implemented under different conditions to commercial banks?	This study will be expected to assist commercial banks in facing competition by dealing effectively with big data.
3	Hartmann and Nair (2010).	How can a demand system for tied goods be applicable in competition? How competing for downstream retail formats are used for customer data analytics? What are the implications of effects for manufacturer pricing when purchasing the tied products through an oligopolistic downstream retail channel in the banking system?	This research has solved competitive challenges by using big data analytic for the banking system.
Profit			
4	Srivastava et al. (2017)	This study has explored how big data analytics can handle this burden of excessive data and many customers? How to use big data analytics to gain profit for organizations in high-level competition? How several issues like fetching tome, connectivity can be handled with big data analytics?	Due to the growing use of the banking system, it has overburdened the existing banking infrastructure. Therefore, this study has solved dealing with excessive customer data without disrupting bank performance and affecting profit.
Crime			
5	Jonker et al. (2012)	How rapid development of analysis application and cyber situational awareness for the 2012 IEEE vast mini-challenge I (mc1)?	This research will provide insight into bank crimes. It has given the solution from the perspective of big data analytics, and it will help reduce bank security issues.
Supply chain			
6	Camerinelli and Schizas (2014).	How to promote supply chain finance at the international level by using big data analytics?	It is expected that this study will be helpful for financiers to develop a system of the supply chain by using big data analytics.
Fintech			
7	Hung and Luo, (2016)	What challenges are faced by banks while investing in fintech companies with the purpose of collaboration?	This research will highlight the strategic consideration for exploring targets for the investment of fintech.
18	Rogers et al. (2016)	What role do fintech companies play in providing services of the supply chain like inventory management and procurement?	In the future, under the light of the study, it is expected that companies of fintech will broaden their operations from financing to supply chain services.
Investment			
9	Biddle et al. (2016)	Does adopting international financial reporting standards (IFRIS) increase capital investment efficiency if measured by value-enhancing risk-taking and investment-cash flow sensitivity?	The study's findings will support IFRISadoptions that can enhance the efficiency of firm-level capital investment, especially in countries with weaker investor protection.
10	Antanavičienė (2014)	What are the approaches to foreign direct investment driving factors and outcomes? What are the current trends of foreign capital flows?	This study will give insights into new consistent patterns of foreign capital directions.

(continued)

Table 9. (continued)

No.	References	Research question	Future expectations
Fraud			
11	Jianhao (2019)	What is the area that is required to be made secure from chances of fraud? What improvement can be made in the anti-fraud system for banks?	This research is to contribute to developing an effective anti-fraud system for reducing bank crimes.
12	Prabowo (2016)	How can security measures be improved in fraud detection in banking by using big data analytics?	This study has found out the kind of algorithms fraud detection that is safe for online banking transactions. Therefore, the research is expected to strengthen the bank security system.
Digital banking			
13	Indriasari et al. (2019)	How can digital banking increase customer satisfaction by using big data analytics?	This research will contribute to the formation of a digital banking system by showing its significance.

and ultimately achieve a competitive advantage. A great example would be the use of big data in determining the aspects of digital banking that increase customer satisfaction using big data analytics to improve that product for profit maximization and customer retention. Evidently, big data is key to sustaining profitability; for example, it allows banks to manage the large volume of existing data without disrupting their operation and profitability. A significant advantage of its implementation in the banking sector is its capability of identifying transactional anomalies or security breaches that alert the banks of the need for security system's improvement or the development of effective anti-fraud systems. The downside, in this case, is the challenges banks face when investing in fintech, such as big data advanced infrastructure.

To the best of the author's knowledge, this objective bibliometric study reveals that there have been only 60 research papers written on big data. It indicates that limited prior research studies on this subject have not received considerable interest from practitioners and researchers. Sivarajah et al. (2017) highlight in their research that big data analytics is widely recognized as a trending practice in companies to construct and derive meaningful information to survive in the market competition. This study further revealed that the number of publishers devoted to this topic has also been limited. Therefore, we recommended that top publishers specialize in identified big data themes, such as digital finance, digital banking, or fintech, on big data subject matters to help the audience comprehend the significance of this big data analytics landscape. Resultantly, robust investment choices by organizations will further contribute toward the discipline of resource allocation and management coupled with the expansion of resources.

This study implemented a bibliometric method to investigate the significance of big data in the banking sector. This investigation gathered articles from the Scopus database, and these articles have been evaluated in VOSviewer software. Sixty shortlisted relevant articles have successfully shed light on the connection between two areas of study: banks

and Big Data. However, this bibliometric analysis has discovered the sluggish growth of researchers' interests in this topic. Limited research on this topic further indicates that this big data concept has emerged recently. These two themes have been combined to explore this topic in different dimensions in the research field. Researchers must fill gaps in this topic by publishing more research papers in comprehensive databases.

This study has further helped us conclude that Scopus is the largest citation database providing peer-reviewed articles compared to Web of Science. Hence, the literature extracted from Scopus for this bibliometric study is considered valuable. Bibliometric analysis in this study has been applied to analyze the latest trends of this research topic. This study has helped in providing evidence of big data analytics significant impact on banks globally. Research collaborators globally on this area of interest have also been identified in this study. This study may further help pave the way for the potential collaboration of researchers in the future on these identified themes.

Moreover, it would enable academic scholars to identify relevant journals to publish their work and reach the right target audience. Most influential journals, countries, and research institutions have been systematically analyzed to study papers on big data. Similarly, the most frequently used keywords in research papers have also been investigated. This paper would be helpful for researchers to track trends and development on big data in banks in this scientific research field.

Studies suggest that big data analytics and applications in the banking sector are numerous. Big data analytics can significantly help banks gauge the risk bank faces when offering risk-associated services to customers, such as loan offerings. Banks are now immersing themselves in digital programs to improve their culture and services. A study by (Xu & Yu, 2019) highlights that big data has brought several changes in many industries across the globe. In their scholarly work, researchers have highlighted the significance of unstructured and structured big data in the corporate sector.

The findings of this study have also helped in understanding the development of big data in financial institutions. It suggests that policymakers in the science field and future researchers can further understand the big data scope in several academic areas. To add on, banks can utilize the information presented in this paper to study and analyze the long-term pros and cons of its implementation in their operation to manage risk, retain clients, build effective models, and so on. Furthermore, big data software developers are to benefit from the findings, as they can clearly comprehend the needs of financial institutions when developing their big data software. In other words, this bibliometric study has reflected the maturity of big data in several domains and emerging current research trends. Mohammadi and Karami (2020) observed that with the flourishing of big data topics globally, other similar themes such as network modeling had received scant attention from researchers.

It is noteworthy that financial institutions, including banks, must function under compliance criteria to gather insights from big data. Banks are initiating to make the best use of derived vital information by introducing software to analyze the big data documents, prevent and detect frauds, and simultaneously improve their standards and quality of services offered to customers. With the banking industry facing immense competition, the integration of big data analytics with the help of innovative techniques is vital to gain optimal output levels. Studies analyzed in this paper have also concluded that big data contributes to risk management, such as credit management, operational risks, fraud management, in banks, and financial institutions. This risk coverage from big data has reduced the response time of banks to customers, thereby enhancing its efficacy levels. Employee engagement is another benefit derived from the valuable information by banks and the financial sector. Researchers have revealed that the best performances of these institutions can be known and acknowledged from analyzing big data. However, researchers have emphasized in their studies that big data can be unpredictable because it varies and changes frequently. Studies have further shown that decisions derived and shaped from big data are more efficient and strategic comparatively. Financial institutions can generate a high return on equities and assets based on data-driven decisions. It infers that the effective use of this dataset can increase the quality of decisions of policymakers in banks. Most importantly, high velocity, volume, and variety of data sets can add value in gathering essential statistics.

This study has applied several bibliometric indicators to analyze publications on this topic in different research realms. A gradual, yet limited, surge in publications on big data in banks reveals the acceptance of its advantages in organizations globally. Moreover, research papers on this domain have been published in differing journals originating from different publishers. Therefore, leading publishers must catalyze research activities and publications on this literature.

Conclusion

With the evolution of technology, significant changes have been made in the banking sector. Big Data analytics is the fastest growing trend because of its influential role in the banking sector. It helps in managing a large amount of data in banks. With the help of a bibliographic review, has analyzed essential themes banks under the scope of big data such as risk assessment, financial management, credit risk, customer analysis, bankruptcy prediction, anti-fraud system, strategic framework management, investment, profit, and competition. The increased importance and benefits of services provided by banks have resulted in increased participation of countries from around the globe to investigate this area of study further.

This research has identified and filtered the documents based on the citation rate and link strength. It proved to be an effective technique in this bibliometric study as the citation rate shows an increased awareness of researchers on this discipline; therefore, they are now more inclined toward shedding light on the significance of this topic by carrying out meaningful research. This paper has mapped the territory and detected new trends on the concept of big data in banks. Evaluative methods have been helped to analyze bibliographic data and to add more valuable insights further. The gathered information can help develop strong customer relations, management decisions, or further information needed in various departments. This information from these quality research papers would outweigh any challenges with big data in time management or storing data.

Furthermore, the studies shed light on extended uses of big data in banks besides customer segmentation. The interest focused on fraud detection, building effective business models, predicting insolvency, and so much more. However, it is established from the resulting number of collected articles that the literature on the topic isn't as developed as other business technologies' literature, such as artificial intelligence's application. Hence, the current literature on big data is limited to primarily the business market, as observed from the analysis. Moreover, it was observed that the main influential topics are based in the United States, China, the United Kingdom, India, and Spain, which further limits the application scope in other countries. The theoretical framework should encompass economic aspects that will either hinder or boost the efficiency of big data in the banking sector of any nation.

Additionally, the authors observed a keen interest in the big data customer analysis-related articles. Still, they didn't delve deeply into, for example, the practical application of big data in fraud detection initiatives. Therefore, the business aspect of those papers, in the context of the baking sector, isn't examined in detail as was expected to form innovative technology papers. Furthermore, non-customer-related collected papers haven't gained much recognition in the research area. It is due to the increased focus on customer satisfaction

and retention to achieve competitiveness and maintain profitability. However, customers aren't the sole profitability predictor, but effective internal operations' management is too highly impactful. Thus, the focus must be shared between all-sustaining business efforts, whether internal or external, and the publication trend should shift to explore other factors that could improve operation and profitability with the application of big data.

Conclusively, big data is a crucial tool in managing massive business data volumes and shouldn't be limited to a single use per bank. On the contrary, banks shifting toward digitization are recommended to use big data to cater valuable products to their clients, manage potential risks, identify fraudulent activities, and build efficient business models.

Limitations

In the conduction of this study, the researchers expected to find more relevant papers on the topic published to incorporate in the analysis of the literature review section. However, they were presented with a minimal number of documents in the banking field on big data within the analyzed period of 60 years. Initially, the investigators were to collect topic-related documents on the Web of Science. Still, through extensive research, it was observed that Scopus is more comprehensive and has a greater citation database of reviewed publications than Web of Science. Thus, the authors had to switch from their original analysis plan and solely conduct a new one based on Scopus.

Future Expectations

This study applied a bibliographic approach to date to evaluate research and available literature on big data in the banking sector. Future studies must further expand this bibliometric analysis scope by using, for example, more documentation, impact factors of the journal, and advanced citation to reveal more insights on this area of study. The research is expected to provide the basis to identify gaps of study for researchers and explore more critical topics of big data in banking comprehensively. Extensive research and theoretical knowledge in this discipline will lead to more achievements in the banking sector and derive more benefits for customers.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iD

Haitham Nobanee  <https://orcid.org/0000-0003-4424-5600>

References

- Abbott, K. W., & Snidal, D. (1998). Why states act through formal international organizations. *Journal of Conflict Resolution*, 42(1), 3–32.
- Al-Dmour, H., Saad, N., Basheer Amin, E., Al-Dmour, R., & Al-Dmour, A. (2021). The influence of the practices of big data analytics applications on bank performance: Filed study. *VINE Journal of Information and Knowledge Management Systems*. Advance online publication. <https://doi.org/10.1108/VJIKMS-08-2020-0151>
- Antanavičienė, J. (2014). Foreign direct investment: Driving factors and outcomes for secure and sustainable development. *Journal of Security and Sustainability Issues*, 3, 55–67.
- Barkin, J. S., & Cronin, B. (1994). The state and the nation: Changing norms and the rules of sovereignty in international relations. *International Organization*, 48(1), 107–130.
- Bhuvana, M., Thirumagal, P. G., & Vasanth, S. (2016). Big data analytics – A leveraging technology for Indian commercial banks. *Indian Journal of Science and Technology*, 9(32), 1–5.
- Biddle, G. C., Callahan, C. M., Hong, H. A., & Knowles, R. (2016). Do adoptions of international financial reporting standards enhance capital investment efficiency? Available at SSRN 2353693.
- Camerinelli, E., & Schizas, E. (2014). *A study of the business case for supply chain finance*. The Association of Chartered Certified Accountants, Studie.
- Chedrawi, C., Atallah, Y., & Osta, S. (2020). Big data in the banking sector from a transactional cost theory (TCT) perspective—the case of top Lebanese banks. *Lecture Notes in Information Systems and Organization*, 35, 391–405.
- Cook, K., Grinstein, G., Whiting, M., & Cooper, M. (2012). *VAST Challenge 2012: Visual analytics for big data* [Conference session]. 2012 IEEE conference on visual analytics science and technology (VAST), Seattle, WA, pp. 251–255. IEEE.
- Dicuonzo, G., Galeone, G., Zappimulso, E., & Dell'Atti, V. (2019). Risk management 4.0: The role of big data analytics in the bank sector. *International Journal of Economics and Financial Issues*, 9(6), 40–47.
- Diniz, E. H., Luvizan, S. S., Hino, M. C., & Ferreira, P. C. (2018). Unveiling the big data adoption in banks: Strategizing the implementation of a new technology. *Lecture Notes in Information Systems and Organization*, 23, 149–162.
- Evangelatos, N., Reumann, M., Lehrach, H., & Brand, A. (2016). Clinical trial Data as public goods: Fair trade and the virtual Knowledge bank as a solution to the free rider problem - a framework for the promotion of iovation by facilitation of clinical trial Data sharing among biopharmaceutical companies in the Era of Omics and Big Data. *Public Health Genomics*, 19(4), 211–219.
- Fahimnia, B., Sarkis, J., & Davarzani, H. (2015). Green supply chain management: A review and bibliometric analysis. *International Journal of Production Economics*, 162, 101–114.
- Fischer, F., Fuchs, J., Mansmann, F., & Keim, D. A. (2015). BANKSAFE: Visual analytics for big data in large-scale computer networks. *Information Visualization*, 14(1), 51–61.
- Fuschi, D. L., & Tvaronavičienė, M. (2014). Sustainable development, big data and supervisory control: Service quality in banking sector. *Journal of Security and Sustainability Issues*, 3, 5–14.

- Guha, A. (2019). Prediction of bankruptcy using big data analytics based on fuzzy C-means algorithm. *IAES International Journal of Artificial Intelligence (IJ-AI)*, 8(2), 168–174.
- Alaka, H., Oyedele, L. O., Bilal, M., Akinade, O. O., Owolabi, H. A., & Ajayi, S. O. (2015). *Bankruptcy prediction of construction businesses: Towards a big data analytics approach* [Conference session]. 2015 IEEE first international conference on big data computing service and applications (pp. 347–352). IEEE.
- Hale, G., & Lopez, J. A. (2019). Monitoring banking system connectedness with big data. *Journal of Econometrics*, 212(1), 203–220.
- Hanrieder, W. F. (1966). International organizations and international systems. *Journal of Conflict Resolution*, 10(3), 297–313.
- Hartmann, W. R., & Nair, H. S. (2010). Retail competition and the dynamics of demand for tied goods. *Marketing Science*, 29(2), 366–386.
- Hassani, H., Huang, X., & Silva, E. (2018a). Digitalisation and big data mining in banking. *Big Data and Cognitive Computing*, 2(3), 18.
- Hassani, H., Huang, X., & Silva, E. (2018b). Banking with blockchain-enabled big data. *Journal of Management Analytics*, 5(4), 256–275.
- Hung, J. L., He, W., & Shen, J. (2020). Big data analytics for supply chain relationship in banking. *Industrial Marketing Management*, 86, 144–153.
- Hung, J. L., & Luo, B. (2016). FinTech in Taiwan: A case study of a bank's strategic planning for an investment in a FinTech company. *Financial Innovation*, 2(1), 1–16.
- Indriasari, E., Gaol, F. L., & Matsuo, T. (2019). *Digital banking transformation: Application of artificial intelligence and big data analytics for leveraging customer experience in the Indonesia banking sector* [Conference session]. 2019 8th international congress on advanced applied informatics (IIAI-AAI), Toyama, Japan, July 7–11, 2019, pp. 863–868. IEEE.
- Jianhao, Y. (2019). Design and implementation of bank wind control anti-fraud project based on big data technology. *Journal of Physics Conference Series*, 1345(2), 022064.
- Jonker, D., Langevin, S., Schretlen, P., & Canfield, C. (2012). *Agile visual analytics for banking cyber "big data"* [Conference session]. 2012 IEEE conference on visual analytics science and technology (VAST), Seattle, WA, October 14–19, 2012, pp. 299–300. IEEE.
- Kalantari, A., Kamsin, A., Kamaruddin, H. S., Ale Ebrahim, N., Gani, A., Ebrahimi, A., & Shamshirband, S. (2017). A bibliometric approach to tracking big data research trends. *Journal of Big Data*, 4(1), 1–18.
- Labrinidis, A., & Jagadish, H. V. (2012). Challenges and opportunities with big data. *Proceedings of the VLDB Endowment*, 5(12), 2032–2033.
- Lee, I. (2017). Big data: Dimensions, evolution, impacts, and challenges. *Business Horizons*, 60(3), 293–303.
- Liu, X., Sun, R., Wang, S., & Wu, Y. J. (2019). The research landscape of big data: A bibliometric analysis. *Library Hi Tech*, 38, 367–384.
- Marín-Marín, J. A., López-Belmonte, J., Fernández-Campoy, J. M., & Romero-Rodríguez, J. M. (2019). Big data in education. A bibliometric review. *Social Science*, 8(8), 223.
- Marr, B. (2016). *Big data in practice: How 45 successful companies used big data analytics to deliver extraordinary results*. John Wiley & Sons.
- Marx, V. (2013). The big challenges of big data. *Nature*, 498(7453), 255–260.
- Mishra, D., Gunasekaran, A., Papadopoulos, T., & Childe, S. J. (2018). Big data and supply chain management: A review and bibliometric analysis. *Annals of Operations Research*, 270(1–2), 313–336.
- Mohammadi, E., & Karami, A. (2020). Exploring research trends in big data across disciplines: A text mining analysis. *Journal of Information Science*. Advance online publication. <https://doi.org/10.1177/0165551520932855>
- Nobanee, H. (2020). Big data in business: A bibliometric analysis of relevant literature. *Big Data*, 8(6), 459–463.
- Nobanee, H. (2021). A bibliometric review of big data in finance. *Big Data*, 9(2), 73–78.
- Nobanee, H., Al Hamadi, F. Y., Abdulaziz, F. A., Abukarsh, L. S., Alqahtani, A. F., AlSubaey, S. K., Alqahtani, S. M., & Almansoori, H. A. (2021). A bibliometric analysis of sustainability and risk management. *Sustainability*, 13(6), 3277.
- Pérez-Martín, A., Pérez-Torregrosa, A., & Vaca, M. (2018). Big data techniques to measure credit banking risk in home equity loans. *Journal of Business Research*, 89, 448–454.
- Prabowo, H. (2016). Learning fraud detection from big data in online banking transactions: A systematic literature review. *Journal of Telecommunication, Electronic and Computer Engineering*, 8(3), 127–131.
- Rahman, N., & Iverson, S. (2015). Big data business intelligence in bank risk analysis. *International Journal of Business Intelligence Research*, 6(2), 55–77.
- Rakhman, R., Widiastuti, R., Legowo, N., & Kaburuan, E. M. (2019). Big data analytics implementation in banking industry – Case study cross selling activity in Indonesia's Commercial bank. *International Journal of Scientific & Technology Research*, 8(9), 1632–1643.
- Reinders, A., Bogatinoska, D. C., Braun, C., Imenes, A. G., Driesse, A., van Slooten, F., Moser, D., Van Sark, W., Oreski, G., Ottersboeck, B., Pearsall, N., Devetakovic, M., & Leloux, J. (2019). *Development of a big data bank for PV monitoring data, analysis, and simulation in COST Action "PEARL PV"* [Conference session]. 2019 IEEE 46th Photovoltaic Specialists Conference, (PVSC), Chicago, IL, June 16–21, 2019, pp.1388–1392. IEEE.
- Reynolds, R. A., Woods, R., & Baker, J. D. (2006). *Handbook of research on electronic surveys and measurements*. IGI Global.
- Rogers, D., Leuschner, R., & Choi, T. Y. (2016). The rise of FinTech in supply chains. *Harvard Business Review*.
- Sagiroglu, S., & Sinanc, D. (2013). *Big data: A review. 2013 international conference on collaboration technologies and systems (CTS)*, San Diego, CA, May 20–24, 2013, pp. 42–47. IEEE.
- Salleh, K. A., & Janczewski, L. (2019). Security considerations in big data solutions adoption: Lessons from a case study on a banking institution. *Procedia Computer Science*, 164, 168–176.
- 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.
- Sobolevsky, S., Massaro, E., Bojic, I., Arias, J. M., & Ratti, C. (2017). *Predicting regional economic indices using big data of individual bank card transactions* [Conference session]. 2017 IEEE international conference on big data, Boston, MA, December 11–14, 2017, pp.1313–1318. IEEE.

- Sobolevsky, S., Sitko, I., Des Combes, R. T., Hawelka, B., Arias, J. M., & Ratti, C. (2014). *Money on the move: Big data of bank card transactions as the new proxy for human mobility patterns and regional delineation. The case of residents and foreign visitors in Spain* [Conference session]. 2014 IEEE international congress on big data, Anchorage, AK, June 27–July 2, 2014 (pp. 136–143). IEEE.
- Srivastava, A., Singh, S. K., Tanwar, S., & Tyagi, S. (2017). *Suitability of big data analytics in Indian banking sector to increase revenue and profitability* [Conference session]. 2017 3rd international conference on advances in computing, communication & automation (ICACCA) (Fall), Dehradun, India, September 15–16, 2017, pp. 1–6. IEEE.
- Srivastava, U., & Gopalkrishnan, S. (2015). Impact of big data analytics on banking sector: Learning for Indian banks. *Procedia Computer Science*, 50, 643–652.
- Sun, N., Morris, J. G., Xu, J., Zhu, X., & Xie, M. (2014). iCARE: A framework for big data-based banking customer analytics. *IBM Journal of Research and Development*, 58(5/6), 4:1–4:9.
- Tang, J., & Karim, K. E. (2019). Financial fraud detection and big data analytics – Implications on auditors' use of fraud brainstorming session. *Managerial Auditing Journal*, 34(3), 324–337.
- Wang, H., Ma, S., & Dai, H. N. (2019). A rhombic dodecahedron topology for human-centric banking big data. *IEEE Transactions on Computational Social Systems*, 6(5), 1095–1105.
- Wang, S., & Petrounias, I. (2017). *Big data analysis on demographic characteristics of Chinese mobile banking users* [Conference session]. 2017 IEEE 19th Conference on Business Informatics (Vol. 2, pp. 47–54). IEEE.
- Wang, Y., Xiuping, S., & Zhang, Q. (2021). Can fintech improve the efficiency of commercial banks?—An analysis based on big data. *Research in International Business and Finance*, 55, 1–28.
- Winther, C., Estabrooks, C. A., Lavis, J. N., Scott, S.D., Wallin, L., & Profetto-McGrath, J. (2009). *A bibliometric analysis of the knowledge utilization literature: The structure and substance of a scientific community technical report (09-01-TR)*. University of Alberta Faculty of Nursing.
- Xian, H., & Madhavan, K. (2014). Anatomy of scholarly collaboration in engineering education: A big-data bibliometric analysis. *Journal of Engineering Education*, 103(3), 486–514.
- Xu, Z., & Yu, D. (2019). A bibliometrics analysis on big data research (2009–2018). *Journal of Data, Information and Management*, 1(1–2), 3–15.
- Yong-Hak, J. (2013). *Web of Science*. Thomson Reuters.
- Zhao, Y., Li, D., & Pan, L. (2015). Cooperation or competition: An evolutionary game study between commercial banks and big data-based e-commerce financial institutions in China. *Discrete Dynamics in Nature and Society*, 2015, 1–8.