Received: 22 Sept 2024 | Revised: 18 Dec 2024 | Accepted: 27 Dec 2024 | Published: 31 Dec 2024
Volume 2(2) Pages 123-147 DOI: https://doi.org/ 10.56279/jicts.v2i2.118

JICTS
Journal of ICT Systems

A Systematic Review of Big Data Techniques, Opportunities, and Challenges for Developing Countries: The Case of Social Media Networks Mining and Analytics

Matendo Didas[1], Shuubi Alphonce Mutajwaa

Center for Information and Communication Technology, St. John's University of Tanzania, Dodoma, Tanzania

[1]Corresponding author
Email: mdidas@sjut.ac.tz

Funding information

This work was funded by St. John's University of Tanzania.

Keywords

Big data
Big data analytics
Big data mining
Predictive analytics
Social media networks

Abstract

Social media networks (SMNs) serve as global communication platforms where users can share content, images, and videos as well as post comments, follow friends, and share their thoughts. However, developing countries are lagging behind in understanding the techniques, challenges, and opportunities associated with mining and analytics of SMNs Big Data (BD). The study's objective was to review relevant literature to establish awareness and understanding in developing countries about these techniques, opportunities, and challenges associated with mining and analytics of SMNs BD. A systematic literature review analysis was used to address the study objective. The SMNs BD mining and analytics techniques resulting from the review include, but are not limited to, data mining, value chain technique, infosphere big insights, and SMNs BD sentiment analysis. Three categories of challenges discovered on the subject under investigation are process challenges, data challenges, management challenges, and infrastructure challenges. Opportunities discovered during the review include, but are not limited to, business improvements and adjustments, constructing intelligent networks, and customer engagement boosts, among others. Based on the review results, the study proposed SMNs BD management, mining, and analytics steps to guide developing countries in any endeavors aiming at utilizing SMNs BD mining and analytics initiatives.

1. Introduction

Social media networks (SMNs) are online communities that are dispersed over a large network of computers [1]. SMNs are used by millions of users worldwide to publish data content in the form of images and videos, change their status, and leave daily comments [2-5]. With the exponential rise in popularity of SMNs, such as Facebook, MySpace, media sharing networks, Tumbler, Twitter, Instagram, YouTube, and micro-blogging, researchers have been encouraged to examine published data content on these platforms [6-8]. The developing countries need to be aware that the main component of SMNs is its information, and they need a lot of it for their ability to persuade people to contribute data about every waking minute [9]. Priceless knowledge about what people and society value is at the core of the relationships formed by the large pool of knowledge held by SMNs firms [9, 10]. In this context, this large amount of information from these SMNs has been described as Big Data (BD), which is referred to as data that is big enough for traditional data tools to process [10-14].

Developing countries need to understand that this massive information from SMNs in a range of sectors can be used to see beyond the mechanics of interaction on how the content contained inside the interactions can affect people's perceptions and company performance. This can only be achieved by raising awareness of mining and analytics of SMNs BD in developing countries [15-18]. Similar to other industries, given proper mining and analytics of SMNs BD, analyzing markets and predicting consumer behavior can be simplified [19, 20].

When it comes to SMNs BD mining and analytics usage, developing countries are still behind compared to developed economies [21, 22]. The number of real-time information streams and people using SMNs is growing rapidly in developing countries [22].

It should be noted by developing countries that tracking trends in SMNs BD during the analytics and mining process can further provide insights on emerging concerns that can be highly relevant to both themselves and global development [23-25]. The existing problem is that developing countries lack systematic understanding and awareness of the potential that mining and analyzing SMNs BD can bring to their private and public service delivery [26].

This paper offers a systematic review in response to this challenge by analyzing the current practices to highlight SMNs BD mining and analytics techniques, obstacles, and opportunities for the developing countries' communities. The following study objective served as the study's guidance to achieve the study intention:

To review the existing potential techniques, opportunities, and challenges associated with SMNs BD mining and analytics for raising awareness and promoting utilization in developing countries.

The remainder of the paper is organized as follows: Section 2 highlights the literature review of the study. Section 3 provides the study's methodology. Section 4 offers a description of the findings and a discussion of the study. The final section offers the study's conclusion and future directions of the study.

2. Method

2.1 SMNs Statistics and SMNs BD Status

SMNs are an effective tool for research, data mining, and analytics due to the large volume of customer data they contain. A significant percentage of BD is sent via SMN content in an unstructured style and the value/ integrity of the

data is guaranteed by consumer-generated SMNsBD [27, 28]. This increases the appeal of SMNs BD to academics, industry, government, and other stakeholders across the globe. Application Programming Interfaces (APIs) are provided by most SMNs, facilitating public and easy access to vast quantities of SMNs BD for research purposes [29]. Numerous SMNs, such as Facebook, Google, Instagram, LinkedIn, and many more, provide businesses and people with unique APIs [29, 30]. A recent brand watch analysis states that out of 7.7 billion people on the planet, 3.499 billion are active SMNs users [31]. This represents 45.4% of the total population and nearly 80% of netizens. The publication date of this report is June 13, 2019, because, between April 2018 and April 2019, there were 202 million more users using SMNs. Approximately 81% of young people think that SMNs are good for one's personal life and the average daily time spent on SMNs by each of these users was 142 minutes. According to this figure, 74 billion dollars were spent by corporations and retailers on SMNs advertising in 2018. To maintain their competitiveness and make the most of their marketing efforts, 91% of retail businesses and 81% of small and medium businesses use SMNs.

Today, users' data has become so revolutionary in volume that it has become the most significant source of huge data. Table 1 displays the amount of involvement on various SMNs platforms (brand Watch's statistics) [31, 32]. With 2.3 billion active users each month, Facebook is the most popular social media network. On YouTube, a platform where users share videos, there were 1.9 billion active users. Similarly, 1.6 billion people use WhatsApp. There were 610 million members on LinkedIn and 330 million active users on Twitter. According to Table 1 [31-34], Facebook has the greatest number of users (2,375 billion) who are also engaged in the platform. Users are also the most active on YouTube (19%), WhatsApp (16%), Instagram (16%), and WeChat (16%).

Table 1. Well-known SMNs and their corresponding monthly users.

S/N	SMNs	Usage
1	Myspace	15 million
2	Facebook	2.375 billion
3	LinkedIn	610 million
4	Periscope	10 million
5	YouTube	1.9 billion
6	WeChat	1 billion
7	Instagram	1 billion
8	WhatsApp	1.6 billion
9	Pinterest	265 million
10	Twitter	330 million
11	Weibo	600 million
12	ResearchGate	17 million
13	Flickr	90 million
14	4chan	22 million
15	Snapchat	190 million
16	Google+	111 million
17	Airbnb	150 million

A vast number of people in SMNs produce tremendous volumes of data every day and this data can be of any format, including structured semi-structured, and unstructured [33, 34]. BD comes from a lot of organized, semi-structured, and mostly unstructured data sources. Because most BD comes from SMNs, it is commonly called "Social Media Networks BD (SMNs BD)" [35, 36]. SMNs BD is a colloquial term that describes the BD derived from and associated with SMNs. The SMNs BD are pertinent and informative to look at and study to make better decisions regarding research, business, and administrative structure. There are two types of SMNs BD: "Social Graph" and "Social Text" [37].

A different group described customer chats on SMNs as "conversations" and the activities taken inside them as "interactions." Aside from its type, social data are disorganized and disorderly, complicating social ties, such as closeness and support among users [38]. The intricate tree-based relationship on several data points makes it difficult for traditional Relational Database Management Systems (RDBMSs) to analyze social data.

RDBMSs and other conventional methods are unable to analyze this SMNs BD. Limited amounts of structured data can be processed using database management systems, or RDBMSs [39, 40].

2.2 The Term "Analytics"

Analytics is the systematic use of statistical techniques in the computation and interpretation of data. To assist with data analysis in formats that are challenging to comprehend, analytics uses statistics, mathematics, and Artificial Intelligence (AI) to help make better judgments. The term "analytics" is also associated with facilitating data analysis by helping to identify patterns, trends, and other insights from unstructured social data [41, 42]. The words "social media networks mining and analytics" and "BD analytics (BDA)" are used interchangeably in this study. Among the various methods of BDA in SMNs are text mining, social graph theory, opinion mining, sentiment analysis, social influence analysis, statistical analysis, cyber risk analysis, and others [43, 44].

2.3 Related studies on SMNs BD

In the subject of BD, there have been numerous reviews thus far and, nevertheless, there are certain shortcomings in the literature evaluations done on this topic. According to systematic research, the study examined the similarities and differences between the present reviews on this subject. Related works are compiled into surveys and systematic literature reviews (SLRs) in Table 4.

For instance, the probable uses of Information Fusion (IF) in SMNs were surveyed by [45]. They also talked about the similarities and variations in SMNs BD processing technologies according to pertinent parameters with a discussion of future research directions and the difficulties in using IF for SMNs BD. According to the findings, applying fusion improves accuracy, dependability, and confidence, but the study made no mention of business intelligence, integration, sharing, security, or data sharing. Furthermore, neither a systematic structure nor a clear indication of the study selection procedure was mentioned in this study.

Additionally, the metadata for BD-related papers found in the Scopus database between 1957 and 2017 was examined [46]. The authors discovered that, given the enormous volume of publications, the actual tendencies in the scholarly BD literature were insufficient for creating real-time indicators. This study was unsystematic, and there were gaps in its privacy, transparency, BD diffusion, and BD quality metrics. Also, publications that had just been published in 2018 or 2019 were not taken into account.

In a different study [47], a review of Social Network Analytics (SNA) was offered, factors such as data sources, features, computational intelligence, analytic approaches, and features of quality were published between 2011 and 2017. Descriptive, diagnostic, predictive, and prescriptive analytics were used to summarize the qualities of BDA for SNA. The BD analytical methodologies were categorized by the authors into text mining, sentiment analysis, modeling, and SNA. The authors grouped the studies based on approaches, methodology, and qualitative characteristics. They gave a thorough overview of the SMNs BD study areas, but the paper selection process was not discussed. However, several issues, including data availability, velocity, quality, and localization, and natural language processing, were left unanswered.

Many more researchers recite several SMNs BD papers. For example, Bukovina [48] examined SMNs technical analysis to study capital market behavior; Martin and Schuurman [49] surveyed SMNs data for qualitative geographic analysis; Arnaboldi et al. [50] investigated the connection between SMNs BD analysis and the accounting function; other authors [51-53] examined BD mining in SMNs. Lastly, further SMNs BD SLRs are carried out, such as the one to use machine

learning (ML) techniques to detect cyber-attacks [55] and the review of firm-level innovations based on text-mining and SMNs analysis [55].

Additionally, it should be mentioned that social influence analysis looks for ways to quantify each user's impact on social networks to determine which individuals are the most influential. New insights into the interactions between web users can be gained through social network analysis. Understanding how users influence one another might help one better understand how information spreads throughout networks.

A thorough examination of social influence analysis is given by Peng and Wang [56], who also examine the features and architecture of social influence analysis using data from extensive social networks. According to these authors, social influence analysis has a great deal of social impact and practical value. It helps us understand people's social behaviors and encourages communication about political, economic, and cultural activities.

Zhang et al. [57] presented a systematic investigation of the effects of online users' reposting activities in a well-known online social networks and users' offline behaviors. The study offers a voting method that determines which users on mobile cellular networks have the greatest influence over the spread of information. By monitoring user interactions, the backbone network's traffic demand was forecast.

Using SMNs BD taken from domestic online news sites, social networking sites, and forums, Zhang et al. [57] and Song et al. [58] investigated and forecasted the trends and patterns associated with youth sexting in South Korea. Moreover, predicting social issues more accurately is made possible by SMNs BD's ability to collect information from a wide range of individuals and combine a vast amount of data.

Taking into account the reviewed publications, the following summary of shortcomings of these studies have been noted: Certain studies have not made explicit reference to the time frames for the reviewed papers. In addition to discussing the study's scope and the time range of the articles, our current paper also takes into account recently published articles. In the existing works, the selection process was not clearly defined because the related papers lacked a systematic construction, and some of the papers were either improperly classified or lacked a taxonomy. Nevertheless, whereas some research has not examined the assessment parameters and evaluation instruments, this paper establishes a subclass for each of them in addition to offering a clear and visual classification.

3. Method

The study employed the Systematic Literature Review (SLR) method to address the research objective. The SLR chosen in this study followed the five phases (Figure 1 and Table 2) [70, 71]. These included coming up with review questions, finding relevant literature, assessing the caliber and quality of the studies considered, gathering and summarizing data, and interpreting the findings.

The exclusion and inclusion criteria

The inclusion and exclusion criteria used in the SLR, outlined in Table 3, were vital for enhancing its scientific value. Papers not published in English or still under publication were excluded. The study aimed to explore approaches, opportunities, and challenges for developing countries in mining and analyzing large social media datasets. To this end, qualitative, quantitative, and case study analyses were included. Additionally, diverse epistemological perspectives were adopted to highlight the interdisciplinary nature of the topic, ensuring a well-rounded exploration of the subject.

Table 2. Five steps of SLR used in the study.

S/N	Step	Description
01	Framing questions for a review	Questions that were clear, concise, and well-organized were used to establish the topics that needed to be covered before the review process began.
02	Identifying relevant work	A comprehensive search for studies using the terms "BDA social media, social media networks BDA, and "data mining" turned up a large number of printed and digital materials. Of the 87 publications that were used in this study, 74 were sampled and selected for a more in-depth examination.
03	Assessing the quality of studies	74 studies were chosen from Step 2 and subsequently submitted to a more comprehensive quality assessment utilizing general critical appraisal guidelines and design-based quality checklists. These assessments were used to look at heterogeneity and offer advice on whether or not a meta-analysis was necessary
04	Summarizing the evidence	Involved data synthesis of tabulating the characteristics of the study.
05	Interpreting the findings	The study interpreted the data and searched for heterogeneity to ascertain whether the overall summary may be believed. Conclusions were made, when appropriate, from the results of high-caliber studies. The benefits and drawbacks of the evidence were used to rank each recommendation.

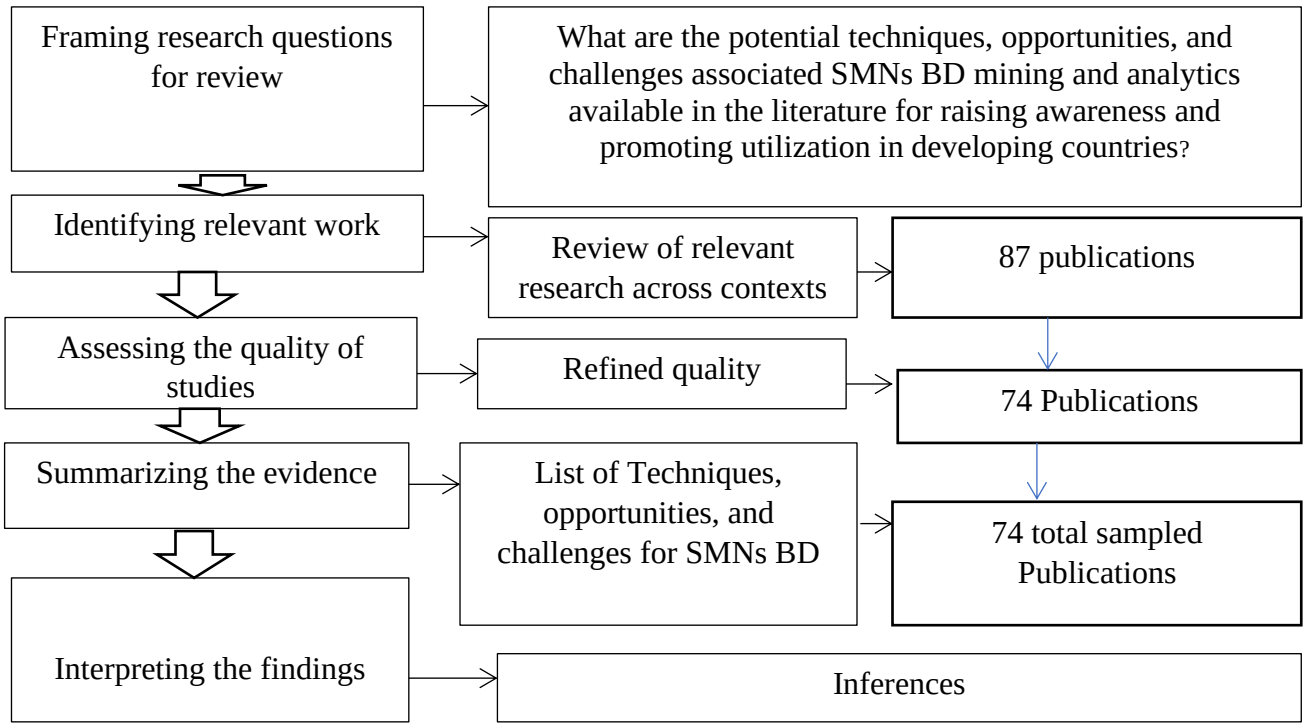

Figure 1. Visual representation of the review stages.

Table 3. The inclusion and exclusion criteria.

Inclusion criteria	Exclusion criteria
Printed and computerized sources, books, Empirical studies, Editorials, Fields of social science, Information science, Economic science, Humanities, English language articles, Articles in academic journals, Case studies, quantitative and qualitative analyses, and epistemological approaches	Summaries of Conferences, Convention lawsuits, Book reviews, Field of agriculture science, Interviews, Technical as well as health science, Summaries of meetings, Editorial letters; non-academic texts, and non-English papers, Papers that were still in publication

4. Results and Discussion

Distribution of publications by database

Figure 2 shows the searched publications from six scholarly databases before applying exclusion and inclusion criteria: ACM Digital Library, IEEE Database, Taylor & Francis, Science Direct-Elsevier, Wiley Online Library, and Scopus.

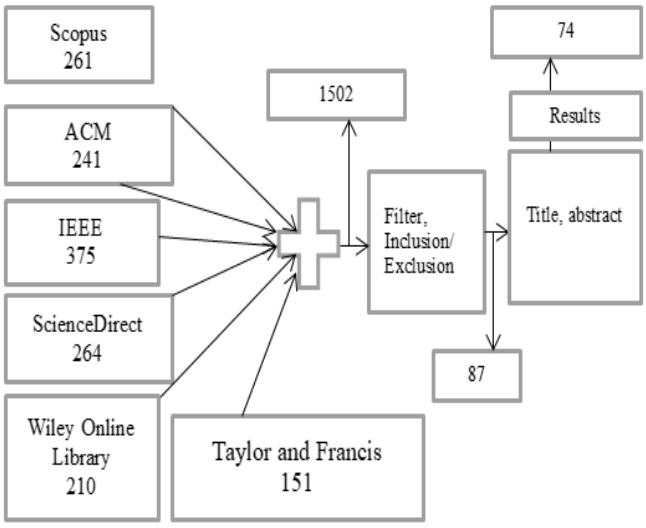

Figure 2. Selection process for review publications.

Figure 3 presents publications selected for analysis by the database. Initially, 87 publications are selected. Following a comprehensive review and evaluation of the abstract and title, the original list was reduced to 74 publications for final analysis.

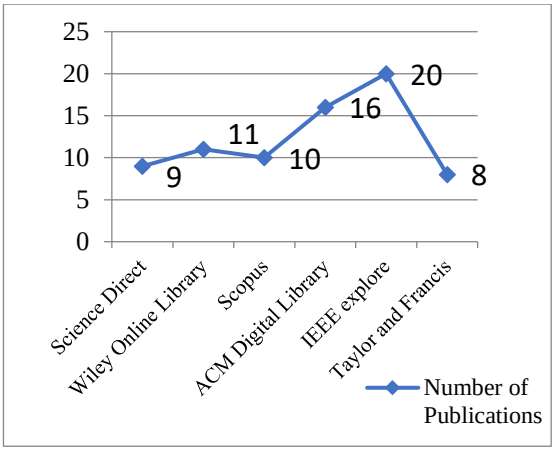

Figure 3. Number of publications by scholarly database source.

Distribution of studies by publication year

Only papers that were published between 2014 and 2024 were picked for the study. It is the best practice for the literature to be recent to capitalize on a time when the related concept, in addition to a commercial dynamic, has found significant resonance. The bulk of the publications were published between 2018 and 2024, indicating the current interest in this field of study (Figure 4).

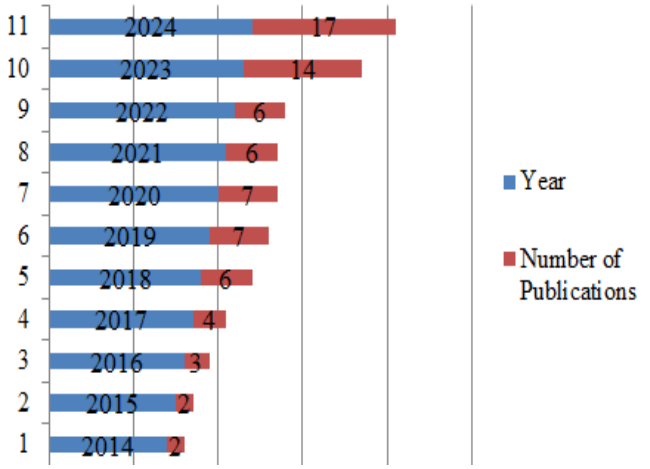

Figure 4: Distribution of studies by publication year.

4.1 SMNs BD mining and analytics techniques for developing countries' understanding and awareness

Developing countries must realize that no single technique can fully capture the vast amount of SMNs BD data. Massive amounts of data can be evaluated using advanced analytics. However, generally, several techniques can be combined to help extract the most value from SMNs BD via mining and analytics processes [72]. The review, however, found a variety of techniques for SMNs BD mining and analysis, with differing outcomes that developing countries need to be aware of, and deployment of the technique relies heavily on the user's level of SMNs BD requirements.

Prediction based on trend detection

This presents an improved technique for prediction backed by trend identification [73, 74]. At first, a dynamic time distortion distance-based K-medoids rule is used to categorize the quality progression of cluster programs into four trends [75, 76]. Then, four trend-specific prediction models are constructed, one at a time, using the random forest regression. By merging the categorization probability with prediction values from the trend-specific models, the intended strategy ultimately produces better prediction outcomes from the SMNs BD [77, 78].

Value chain

This technique emphasizes dividing SMNs BD during the mining and analytics process into four different models (information generation, collection, storage, and analytics). Together, these four elements comprise a huge information value chain [79]. The popular technologies that can be used under this technique include the Hadoop Distributed File system (HDFS) [79].

Data mining

The increasing reliance on SMNs BD necessitates the use of data processing techniques for mining and analytics to arrange unstructured data in a way that follows a scientific pattern [80]. The technique is concerned with examining the information mining class of methods for SMNs BD platforms [81, 82].

InfoSphere big insights

This targets SMNs BD usage across a range of planned subjects [83, 84]. It follows a detailed description for processing data, which is mostly collected from SMNs by Apace Flume and stored in Hadoop storage [84].

SMNs BD sentiment analysis

It presents a revolutionary technique for managing large volumes of SMNs BD during the mining and analytics process [85]. It includes a generative framework for the philosophies of procedure science “theory of social information, abstract and formal models of social information, and an analytical framework for combining large-scale social information sets with the structure and unstructured social information sets [85]”. Table 4 presents a review of other diminishing SMNs BD mining and analytics techniques.

4.2 SMNs BD mining and analytics opportunities for developing countries' understanding and awareness

In the previous sub-section, we reviewed the SMNs BD mining and analytical techniques for understanding and awareness by developing countries. However, applying the above-reviewed techniques to SMNs BD mining and analytics can provide tremendous value in various sectors of developing countries. In this sub-section, the review provides several major opportunities for mining and analyzing SMNs BD in the interest of developing countries' understanding and awareness.

Table 4. List of diminishing data mining techniques for SMNs BD.

Technique	Description	Reference
Structural equivalencemeasures	Uses common behavior to determine the friendship structure of SMNs.	[86]
Friend of a Friend	Used to investigate the growth and evolution of groups at the local and global community levels within extensive social networks on the Semantic Web.	[87]
Semantic Web-based Social Network Analysis Model	To achieve intelligent retrieval of the Web services, the ontological field library of social network analysis is created by combining it with the traditional semantic web outline.	[88]
VoyeurServer	Enhanced the open-source Web-Harvest framework for gathering data from online social networks to investigate online scientific associations and trust-enhancing structures.	[89]
Support Vector Machine/ linear kernel	Used to determine the polarity of documents' neutral instances.	[90]
Social-Unionmethod	Utilized to take advantage of multi-modal SMNs that propose items.	[91]
Document-Pivot Topic Detection, Graph-based Feature-Pivot Topic Detection, Latent Dirichlet Allocation, Frequent Pattern Mining, and Soft Frequent Pattern Mining	Used to detect real-world events on SMNs data	[92]
First StoryDetection	Used to use real-time indications from SMNs data streams to identify both expected and unpredictable events.	[93]
Latent Dirichlet Allocation topic inference model	Used to determine unexpected spikes in the mention of particular hash tags.	[94]
Transaction-basedRule Change Mining	Used to connect changes in real-world news and events with changes in the Association Rules found in tweet hash tags.	[95]

Business improvements and adjustment

For large amounts of data, SMNs collectively are the most important source of BD and interested parties can compile and review all actions that users take on any SMNs, including likes, tweets, views, comments, and favorites for their business improvements and adjustment [96].

Constructing intelligent networks

SMN BD on social transportation has created previously unheard-of chances to construct the intelligent transportation networks of the future. Zheng et al. [97] examined several strategies, such as data sources, analytical techniques, and application systems that are required for using SMNs BD in social transportation systems. For instance, a real-time monitoring system for traffic incident detection via Twitter stream analysis was proposed by D'Andrea et al. [98]. Also, Zhou et al. [99] investigated the role that real-world SMNs BD can play in developing an intelligent transportation sector.

Customer engagement boost

In the digital age, SMNs BD mining and analytics are essential for any business to maintain an online presence, since it enables customer engagement on a personal level [100]. Hence, it causes numerous advantages in business value creativity and productivity.

Creating new opportunities

SMNs BD mining and analytics can help organizations in developing countries make better use of their data to find new opportunities [101]. For example, Facebook alone has two billion monthly users, which make up roughly 26% of the world's population. It is vital to keep in mind that a plethora of information from SMNs will come in a bewildering variety of formats for a variety of informed decision-making processes [102, 103].

Event detection

Mining and analyzing SMNs BD has been used to detect events, such as natural catastrophes [104]. By disseminating emergency information to communities impacted by the crisis, SMNs BD plays a significant role in disaster management. To better understand the characteristics and create relief plans, academics have previously used data from SMNs [104,105]. During the 2016 Louisiana flood, Kim et al. [105] investigated patterns created by online users' interactions on Facebook. Yoo et al. [106] used data from Twitter during Hurricane Sandy to assess the factors that influence the pace at which information spreads. For disaster management, Choi and Bae [107] present Social Big Board, a real-time social BD monitoring system.

Profitable business choices

SMNs BD is the precious oil today, SMNs being a robust source of BD nowadays, if utilized through mining and analytics processes can lead to more profitable business choices, more economical operations, satisfied customers, and more earnings [108].

Cost reduction

When massive amounts of knowledge from SMNs BD are stored, large-scale information technologies like Hadoop and cloud-based analytics can reveal more economical methods to do business and will produce major cost reductions [109, 110].

SMNs BD in economics

Economic growth has been predicted using SMNs textual data. By examining large amounts of data from SMNs blogs, Yamada et al. [111] presented a method to estimate the economic indicator. Their method can drastically cut down on the time lag and yield findings almost instantly, in contrast to the government's announcement of the economic index, which often has a delay because of the time needed to collect and evaluate data. Spagnuolo et al. [112] examined Bitcoin trust-weighted signed network statistics and tracked users' financial circumstances using associated identities to keep an eye on the Bitcoin economy.

Fast-order thinking with new services, and products

SMNs BD mining and analytics can also lead to quicker, higher-order thinking [113]. With HDFS's speed and in-memory analytics, businesses can now evaluate data and make decisions based on what they have learned along with exploring new knowledge sources.

4.3 SMNs BD Mining and Analytics challenges for developing countries' understanding and awareness

Though ironic content on SMNs BD mining and analytics brings unprecedented opportunities, it also comes with various challenges. In the remainder of this section, we summarize the key

challenges under this sub-section associated with SMNs BD mining and analytics for developing countries' understanding and awareness. There are limited studies on the potential challenges of SMNs BD mining and analytics in poor countries [114-116].

The following succinctly describes the challenges associated with SMNs BD mining and analytics for understanding and awareness by developing countries, given their desire to adapt mining and analytics of SMNs BD [117, 118]. The review grouped the challenges into the following categories (Table 5).

Data challenges

Data transfer, privacy violations, sharing, sharing, analyzing, storing, searching, and sharing are among the data challenges [119]. Database systems from the past are also insufficient, and designing the perfect architecture for SMNs mining and analytics that manages historical and real-time data at the same time poses issues [119, 120].

Process challenges

These are related to the group of "how" methods, which include how to gather, combine, process, select the suitable model, and technique for analysis, and display the outcomes during SMNs BD mining and analytics process [121-124]. Consequently, these challenges can be further subdivided into the data value chain stages. Data acquisition scenarios frequently involve large volume, high velocity, high diversity, and low-value data [125-127]. Understanding varied computational difficulties, information security, and computational ways to analyze BD is required to tackle the problems [128-130]. A lack of understanding of the demographics and decision-making processes of SMNs BD users can make it difficult to interpret and process the content of SMNs posts with confidence [131, 132].

Management challenges

These are concerned with the techniques and tools needed to handle data effectively and derive meaningful conclusions from large, varied SMNs BD set that support organizational decision-making [133]. Because SMNs BD comprises a large amount of complex data, it can be difficult for an organization to classify the data by privacy levels and put security measures in place by these levels [134]. Beyond volume and variety of sources, SMNs BD incorporates characteristics, such as scale, speed, structure, and quality [135]. The six types of data management challenges, namely privacy, security, governance, sharing of data and information, cost/operational expenses, and data ownership have been delineated and reviewed [135-139]. Data protection management from security breaches should be an organization's top priority as far as SMNs BD is concerned [140-144].

During the management of SMNs BD mining and analytics process, it is also beneficial to examine regulations, policies, standards, decision-making power, accountability, and enforcement tactics, and data and information ownership and sharing [145, 146]. However, when large-scale datasets are stored, communicating and integrating critical information inside an organization or between various companies can be a difficult undertaking. Sharing data and information will help organizations coordinate with other stakeholders and develop strong relationships. Therefore, it is critical to find a balance and exercise control over the necessity of sharing for optimum impact.

Infrastructure challenges

This involves Limited technological infrastructure for executing SMNs BD mining and analytics process.

Table 5. SMNs BD mining and analytics challenges for developing countries understanding and awareness.

Challenge category and its breakdown	Context	Source
Data challenges		
Designing an analytics architecture to cater to 7 Vs with legacy databases	SMNs BD	[147]
Lack of data expertise	SMNs BD	[148]
Lack of reliable data sources for data collection	SMNs BD	[148]
Infrastructure		
Limited technological infrastructure in developing countries	SMNs BD	[149]
Process		
Lack of value to service delivery due to unreliable alignment between social media initiatives and government strategies	Service Delivery to citizens; social media; Government	[150]
Lack of fault tolerance techniques	SMNs BD mining and analytics	[151]
Visualization of data	SMNs BD	[152, 153]
Management challenges		
Data privacy, security, and control issues due to heterogeneous data and data sources	SMNs BD	[154]
The lack of SMNs BD management presents difficulties for the government to sort this data on privacy levels and to apply security according to these levels	SMNs BD; Government	[155, 156]
The lack of SMNs BD infrastructure has compromised the security, privacy & confidentiality of data through unintended, unauthorized access, or inappropriate access by privileged users	SMNs BD	[156]
The lack of data management tools and techniques results in a negative impact on the decision-making process of the government	Government; Decision making	[157,158]
The lack of SMNs BD governance causes low levels of accessibility for social media BD mining and analytics	SMNs BD mining and analytics	[159]
The lack of data ownership results in the quick spread of incorrect or false information	SMNs BD	[160]
Operational costs & budget allocations	Decision making	[161]
A lack of skills for SMA & related tools presents difficulties for the government to interpret data	SMNs BD mining and analytics	[162]

Figure 5 deals with the steps related to the SMNs BD management. From the review literature, the study proposes a) the steps to be taken by developing countries in managing SMNs BD (Figure 6) and b) the summary of the associations between the reviewed challenges of SMNs BD mining and analytics for developing countries' awareness (Figure 5).

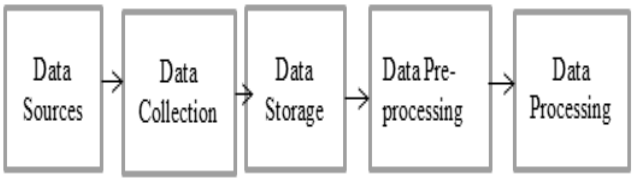

Figure 5. SMNs BD management steps.

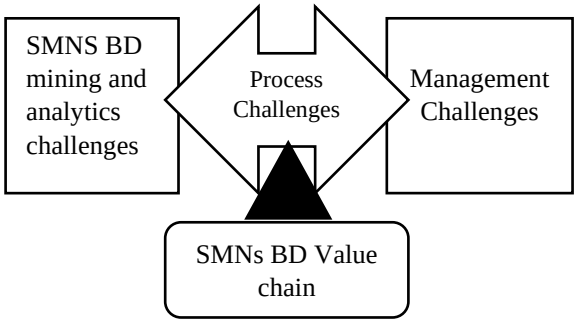

Figure 6. Study's observation framework for SMNs BD mining and analytics challenges for developing countries' awareness.

The overall conclusion is that the phrase "BD" is highly popular right now. Businesses and organizations are discussing their BD solutions and analytical applications. The data used in these applications come from a variety of sources. However, developing nations must understand that SMNs BD is a sort of data that most corporations are very interested in and that mining and analytics of this data can have positive social effects.

Successful mining and analysis of this data can be achieved with a high return on investment as long as developing nations are aware of the methods, opportunities, and difficulties related to SMNs mining and analytics. A significant portion of the global population uses SMN's BD applications. An essential component of modern civilization is the capacity to communicate instantaneously and connect with others and businesses across great distances. Users can exchange media, thoughts, opinions, and comments with friends, family, companies, and organizations via SMN's BD applications. Many different kinds of organizations can benefit from the information found in these remarks, concepts, and mediums. It is feasible to forecast particular application user behavior through data mining and analytics.

5. Conclusion

The study's goal was to review the potential techniques, obstacles, and opportunities associated with SMNs BD mining and analytics for developing countries' understanding and awareness. The study reviewed SMNs BD mining and analytics techniques, such as data mining, value chain technique, infosphere big insights, and SMNs BD sentiment analysis. Three categories of challenges discovered on the subject under investigation are process challenges, data challenges, management challenges, and infrastructure challenges. On the other hand, opportunities discovered during the review include, but are not, limited to business improvements and adjustments, constructing intelligent networks, and customer engagement boosts among others. Based on the review results, the study proposed SMNs BD management, mining, and analytics steps to guide developing countries in any endeavors aiming at utilizing SMNs BD mining and analytics projects. The study also envisions the SMNs BD management steps that developing countries can leverage towards their endeavors and initiatives towards SMNs BD mining and analytics. Because the study is based on secondary data from previous studies, one of its limitations is that it lacks an empirical evaluation in a practical setting. Further works can look into the empirical application of the reviewed thematic areas. To summarize, mining and analytics from SMNs BD have the potential to

contribute to the timely advancement of research and development in a variety of disciplines for developing countries.

CONTRIBUTIONS OF CO-AUTHORS

Matendo Didas	Conceived the idea wrote the paper and finalized the paper.
Shuubi A. Mutajwaa	Provided the technical information of the paper as well as reviewed and proofread the initial draft.

REFERENCES

[1] U. Can and B. Alatas, "A new direction in social network analysis: Online Social Network Analysis Problems and applications," *Physica A: Statistical Mechanics and its Applications*, vol. 535, pp. 122372, 2019.

[2] A. Arora, S. Bansal, C. Kandpal, R. Aswani, and Y. Dwivedi, "Measuring social media influencer index- insights from Facebook, Twitter and Instagram," *Journal of Retailing and Consumer Services*, vol. 49, pp. 86–101, 2019.

[3] W. K. Lai, Y. U. Chen, and T.-Y. Wu, "Analysis and evaluation of random-based message propagation models on the Social Networks," *Computer Networks*, vol. 170, p. 107047, 2020.

[4] A. A. Alalwan, N. P. Rana, Y. K. Dwivedi, and R. Algharabat, "Social Media in Marketing: A review and analysis of the existing literature," *Telematics and Informatics*, vol. 34, no. 7, pp. 1177–1190, 2017.

[5] R. Kumar, J. Novak, and A. Tomkins, "Structure and evolution of online social networks," *Proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining*, pp. 611–617, 2006.

[6] Y. Feng, P. Zhou, D. Wu, and Y. Hu, "Accurate content push for content-centric social networks: A Big Data Support Online Learning Approach," *IEEE Transactions on Emerging Topics in Computational Intelligence*, vol. 2, no. 6, pp. 426–438, 2018.

[7] J. Heidemann, M. Klier, and F. Probst, "Online social networks: A survey of a global phenomenon," *Computer Networks*, vol. 56, no. 18, pp. 3866–3878, 2012.

[8] A. H. Busalim and A. R. C. Hussin, "Understanding social commerce: A systematic literature review and directions for further research," *Int. J. Inf. Manage.*, vol. 36, no. 6, pp. 1075–1088, 2016.

[9] B. Singh, N. F. Raun, and N. A. Sole, "Contemporary social media and IoT-based pandemic control: Exploring possibilities of big data analytics for healthcare governance," in *Big Data Analytics in Cognitive social media and Literary Texts*, Singapore: Springer Singapore, 2021, pp. 101–117.

[10] N. A. Ghani, S. Hamid, I. A. Targio Hashem, and E. Ahmed, "Social media big data analytics: A survey," *Comput. Human Behav.*, vol. 101, pp. 417–428, 2019.

[11] A. A. Osho, "Privacy and security implications of active participation in online social networks: An information diffusion-based approach to modeling user behavioral patterns", PhD Dissertation, Kansas State University, 2022.

[12] M. S. Hadi, A. Q. Lawey, T. E. H. El-Gorashi, and J. M. H. Elmirghani, "Big data analytics for wireless and wired network design: A survey," *Comput. Netw.*, vol. 132, pp. 180–199, 2018.

[13] A. Gandomi and M. Haider, "Beyond the hype: Big data concepts, methods, and analytics," *Int. J. Inf. Manage.*, vol. 35, no. 2, pp. 137–144, 2015.

[14] R. Kitchin, "Big Data, new epistemologies and paradigm shifts," *Big Data Soc.*, vol. 1, no. 1, p. 205395171452848, 2014.

[15] S. Sagiroglu and D. Sinanc, "Big data: A review," in *2013 International Conference on Collaboration Technologies and Systems (CTS)*, 2013.

[16] F.-Q. Pei, D.-B. Li, and Y.-F. Tong, "Double-layered big data analytics architecture for solar cells series welding machine," *Comput. Ind.*, vol. 97, pp. 17–23, 2018.

[17] G. Bello-Orgaz, J. J. Jung, and D. Camacho, "Social big data: Recent achievements and new challenges," *Inf. Fusion*, vol. 28, pp. 45–59, 2016.

[18] S. Peng, G. Wang, and D. Xie, "Social influence analysis in social networking big data: Opportunities and challenges," *IEEE Netw.*, vol. 31, no. 1, pp. 11–17, 2017.

[19] S. Matilda, "Big data in social media environment: A business perspective," in *Decision Management: Concepts, methodologies, tools, and applications*, IGI Global, pp. 1876–1899, 2017.

[20] T. S. Landers, *Why and How Local Nonprofit Organizations Use Facebook Compared to*. Other social media, 2024.

[21] J. Poushter, C. Bishop, and H. Chwe, "Social media use continues to rise in developing countries but plateaus across developed ones," *Pew Research Center*, vol. 22, pp. 2–19, 2018.

[22] A. M. Chew and D. V. Gunasekeran, "Social media big data: the good, the bad, and the ugly (un) truths," *Frontiers in Big Data*, vol. 4, 2021.

[23] P. Arora, "Bottom of the data pyramid: Big data and the global south," *International Journal of Communication*, vol. 10, 2016.

[24] S. B. Abkenar, M. H. Kashani, E. Mahdipour, and S. M. Jameii, "Big data analytics meets social media: A systematic review of techniques, open issues, and future directions," *Telematics and informatics*, vol. 57, 2021.

[25] Y. Duan, J. S. Edwards, and Y. K. Dwivedi, "Artificial intelligence for decision making in the era of Big Data-evolution, challenges, and research agenda. International journal of information management," vol. 48, pp. 63–71, 2019.

[26] C. Udanor, S. Aneke, and B. O. Ogbuokiri, "Determining social media impact on the politics of developing countries using social network analytics," *Program*, vol. 50, no. 4, pp. 481–507, 2016.

[27] R. Moro Visconti, A. Larocca, and M. Marconi, "Big data-driven value chains and digital platforms: From value co-creation to monetization," *SSRN Electronic Journal*, 2017.

[28] P. J. H. Daas, M. J. Puts, B. Buelens, and P. A. M. van den Hurk, "Big Data as a source for official statistics," *J. Off. Stat.*, vol. 31, no. 2, pp. 249–262, 2015.

[29] M. J. Piskorski, *A social strategy: How we profit from social media*. Princeton University Press, 2016.

[30] M. M. Maja and P. Letaba, "Towards a data-driven technology roadmap for the bank of the future: Exploring big data analytics to support technology road mapping," *Soc. Sci. Humanit. Open*, vol. 6, no. 1, 2022.

[31] N. J. Time, "The History of Facebook: From Basic to Global Giant, Brandwatch," *Pew Rsch. Ctr*, 2014.

[32] M. Gómez and M. De Luna, "The role of social networks in communication in the scientific research community," *JOTSE: Journal of Technology and Science Education*, vol. 14, no. 2, pp. 291–305, 2024.

[33] D. Camacho, À. Panizo-LLedot, G. Bello-Orgaz, A. Gonzalez-Pardo, and E. Cambria, "The four dimensions of social network analysis: An overview of research methods, applications, and software tools," *arXiv [cs.SI]*, 2020.

[34] A. Sapountzi and K. E. Psannis, "Social networking data analysis tools & challenges," *Future Gener. Comput. Syst.*, vol. 86, pp. 893–913, 2018.

[35] V. Rajaraman, "Big Data Analytics," *Resonance*, vol. 21, no. 8, pp. 695–716, 2016.

[36] K. Lyko, M. Nitzschke, and A.-C. Ngonga Ngomo, "Big Data Acquisition," *New Horizons for a Data-Driven Economy*, pp. 39–61, 2016.

[37] B. E. Lopez, N. R. Magliocca, and A. T. Crooks, "Challenges and opportunities of social media data for socio-environmental systems research," *Land*, vol. 8, no. 7, p. 107, 2019.

[38] M. A.-U.-D. Khan, M. F. Uddin, and N. Gupta, "Seven V's of Big Data understanding Big Data to extract value," in *Proceedings of the 2014 Zone 1 Conference of the American Society for Engineering Education*, 2014.

[39] M. Kvet, J. Papan, and M. H. Durneková, "Treating temporal function references in relational database management system," *IEEE Access*, 2024.

[40] J. J. Pan, J. Wang, and G. Li, "Survey of vector database management systems," *VLDB J.*, vol. 33, no. 5, pp. 1591–1615, 2024.

[41] M. S. Rahman and H. Reza, "A systematic review towards big data analytics in social media," *Big Data Min. Anal.*, vol. 5, no. 3, pp. 228–244, 2022.

[42] P. Brooker, J. Barnett, and T. Cribbin, "Doing social media analytics," *Big Data Soc.*, vol. 3, no. 2, pp. 205395171665806, 2016.

[43] Z. Lei, Y. Chen, and M. K. Lim, "Modeling and analysis of big data platform group adoption behavior based on social network analysis," *Technology in Society*, vol. 65, 2021.

[44] M. N. Noor and F. Haneef, "A review of big data and social network analytics techniques," *Researchpedia Journal of Computing*, vol. 1, no. 1, pp. 39–49, 2020.

[45] I. Yaqoob *et al.*, "WITHDRAWN: Information fusion in social big data: Foundations, state-of-the-art, applications, challenges, and future research directions," *Int. J. Inf. Manage.*, 2016.

[46] D. Bella, E. Leporatti, and L. Maggino, "Big data and social indicators: Actual trends and new perspectives," *Social Indicators Research*, vol. 135, pp. 869–878, 2018.

[47] D. B. Kurka, A. Godoy, and F. J. Von Zuben, "Online social network analysis: A survey of research applications in computer science," *arXiv [cs.SI]*, 2015.

[48] J. Bukovina, "Social media big data and capital markets—An overview," *J. Behav. Exp. Finance*, vol. 11, pp. 18–26, 2016.

[49] M. E. Martin and N. Schuurman, "Social media big data acquisition and analysis for qualitative GIScience: Challenges and opportunities," *Ann. Am. Assoc. Geogr.*, vol. 110, no. 5, pp. 1335–1352, 2020.

[50] H. Esfahani, K. Tavasoli, and A. Jabbarzadeh, "Big data and social media: A scientometrics analysis," *International Journal of Data and Network Science*, vol. 3, no. 3, pp. 145–164, 2019.

[51] I. Guellil and K. Boukhalfa, "Social big data mining: A survey focused on opinion mining and sentiments analysis," in *2015 12th International Symposium on Programming and Systems (ISPS)*, 2015.

[52] S. Gole and B. Tidke, "A survey of big data in social media using data mining techniques," in *2015 International Conference on Advanced Computing and Communication Systems*, 2015.

[53] P. V. Paul, K. Monica, and M. Trishanka, "A survey on big data analytics using social media data," in *2017 Innovations in Power and Advanced Computing Technologies (i-PACT)*, 2017.

[54] M. A. Al-Garadi *et al.*, "Predicting cyberbullying on social media in the big data era using machine learning algorithms: Review of literature and open challenges," *IEEE Access*, vol. 7, pp. 70701–70718, 2019.

[55] O. Lerena, F. Barletta, F. Fiorentin, D. Suárez, and G. Yoguel, "Big data of innovation literature at the firm level: a review based on social network and text mining techniques," *Econ. Innov. New Technol.*, vol. 30, no. 2, pp. 134–150, 2021.

[56] S. Peng, G. Wang, and D. Xie, "Social influence analysis in social networking big data: Opportunities and challenges," *IEEE Netw.*, vol. 31, no. 1, pp. 11–17, 2017.

[57] Y. Zhang *et al.*, "Mobile social big data: WeChat moments dataset, network applications, and opportunities," *IEEE Netw.*, vol. 32, no. 3, pp. 146–153, 2018.

[58] J. Song, T. M. Song, and J. R. Lee, "Stay alert: Forecasting the risks of sexting in Korea using social big data," *Comput. Human Behav.*, vol. 81, pp. 294–302, 2018.

[59] H. Sebei, M. A. Hadj Taieb, and M. Ben Aouicha, "Review of social media analytics process and Big Data pipeline," *Soc. Netw. Anal. Min.*, vol. 8, no. 1, 2018.

[60] A. Subroto and A. Apriyana, "Cyber risk prediction through social media big data analytics and statistical machine learning," *J. Big Data*, vol. 6, no. 1, 2019.

[61] C. Rammer and N. Es-Sadki, "Using big data for generating firm-level innovation indicators-a literature review," *Technological Forecasting and Social Change*, vol. 197, 2023.

[62] J. P. Verma and S. Agrawal, "Big data analytics: Challenges and applications for text, audio, video, and social media data," *Int. J. Soft Comput. Artif. Intell. Appl.*, vol. 5, no. 1, pp. 41–51, 2016.

[63] C. A. Pushpam and J. G. Jayanthi, "Overview of data mining in social media," *International Journal of Computer Sciences and Engineering*, vol. 5, no. 11, pp. 147–157, 2017.

[64] J. Peng, A. Agarwal, K. Hosanagar, and R. Iyengar, "Network overlap and content sharing on social media platforms," *J. Mark. Res.*, vol. 55, no. 4, pp. 571–585, 2018.

[65] B. Abu-Salih, P. Wongthongtham, D. Zhu, K. Y. Chan, and A. Rudra, "Predictive analytics using Social Big Data and machine learning," *arXiv [cs.CY]*, 2021.

[66] M. Arnaboldi, C. Busco, and S. Cuganesan, "Accounting, accountability, social media and Big Data: Revolution or hype?," Accounting, Auditing & Accountability Journal, vol. 30, no. 4, pp. 762–776, May 2017.

[67] M. E. Martin and N. Schuurman, "Social media big data acquisition and analysis for qualitative GIScience: Challenges and opportunities," *Ann. Am. Assoc. Geogr.*, vol. 110, no. 5, pp. 1335–1352, 2020.

[68] D. Vecchio, P. Mele, G. Ndou, and V. Secundo, "Open innovation and social big data for sustainability: Evidence from the tourism industry," *Sustainability*, vol. 10, 2018.

[69] S. Gao, H. Pang, P. Gallinari, J. Guo, and N. Kato, "A novel embedding method for information diffusion prediction in social network big data," *IEEE Trans. Industr. Inform.*, vol. 13, no. 4, pp. 2097–2105, 2017.

[70] K. S. Khan, R. Kunz, J. Kleijnen, and G. Antes, "Five steps to conducting a systematic review," *J. R. Soc. Med.*, vol. 96, no. 3, pp. 118–121, 2003.

[71] J. P. T. Higgins, S. G. Thompson, J. J. Deeks, and D. G. Altman, "Measuring inconsistency in meta-analyses," *BMJ*, vol. 327, no. 7414, pp. 557–560, 2003.

[72] D. Valle and L. Kenett, "Social media big data integration: A new approach based on calibration," *Expert Systems with Applications*, vol. 111, pp. 76–90, 2018.

[73] R. Mohamadrezaei, R. Ravanmehr, "Trend Detection and Prediction in Blogosphere based on Sentiment Analysis using PSO and Q-Learning," *International Journal of Information & Communication Technology Research*, vol. 12, no. 1, pp. 42–55, 2020.

[74] X. Dinh, T. Trinh, T. Ngoc, and V. Pham, *Improving Social Trend Detection Based on User Interaction and Combined with Keyphrase Extraction Using Text Features on Word Graph. InIntelligent Systems and Networks: Selected Articles from ICISN 2021*. Vietnam; Singapore: Springer, 2021.

[75] M. D. Salas-Zárate, J. Medina-Moreira, P. J. Álvarez-Sagubay, K. Lagos-Ortiz, M. A. Paredes-Valverde, and R. Valencia-García, *Sentiment analysis and trend detection in Twitter. InTechnologies and Innovation: Second International Conference, CITI 2016*. Guayaquil, Ecuador: Springer International Publishing, 2016.

[76] H. M. Subrahmanya and T. Shivaprakash, "Development of big data dimensionality reduction methods for effective data transmission and feature enhancement algorithms," Lecture Notes in Electrical Engineering, pp. 1–8, 2023.

[77] C. Zhu, G. Cheng, and K. Wang, "Big data analytics for program popularity prediction in broadcast TV industries," *IEEE Access*, vol. 5, pp. 24593–24601, 2017.

[78] P. Sahatiya, "Big data analytics on social media data: a literature review," *International Research Journal of Engineering and Technology*, vol. 5, no. 2, pp. 189–192, 2018.

[79] H. Hu, Y. Wen, T.-S. Chua, and X. Li, "Toward scalable systems for big data analytics: A technology tutorial," *IEEE Access*, vol. 2, pp. 652–687, 2014.

[80] B. Ravinder *et al.*, "Web data mining with organized contents using naive Bayes algorithm," *2024 2nd International Conference on Computer, Communication and Control (IC4)*, pp. 1–6, Feb. 2024.

[81] M. Injadat, F. Salo, and A. B. Nassif, "Data mining techniques in social media: A survey," *Neurocomputing*, vol. 214, pp. 654–670, 2016.

[82] A. Kumar, S. R. Sangwan, and A. Nayyar, "Multimedia Social Big Data: Mining," *Intelligent Systems Reference Library*, pp. 289–321, 2019.

[83] M. Birjali, A. Beni-Hssane, and M. Erritali, "Analyzing social media through big data using infosphere big insights and Apache flume," *Procedia computer science*, vol. 113, pp. 280–285, 2017.

[84] M. Popović and M. Milosavljević, "Twitter data analytics in education using IBM infosphere biginsights," in *Proceedings of the International Scientific Conference - Sinteza 2016*, 2016.

[85] Q. A. Xu, V. Chang, and C. Jayne, "A systematic review of social media-based sentiment analysis: Emerging trends and challenges," *Decision Analytics Journal*, vol. 3, no. 100073, p. 100073, 2022.

[86] Z. Lin, Y. Zhang, Q. Gong, Y. Chen, A. Oksanen, and A. Y. Ding, "Structural hole theory in social network analysis: A review," *IEEE Trans. Comput. Soc. Syst.*, vol. 9, no. 3, pp. 724–739, 2022.

[87] M. Amith, K. Fujimoto, R. Mauldin, and C. Tao, "Friend of a Friend with Benefits ontology (FOAF+): extending a social network ontology for public health," *BMC Med. Inform. Decis. Mak.*, vol. 20, no. Suppl 10, p. 269, 2020.

[88] M. Barhamgi, A. Masmoudi, R. Lara-Cabrera, and D. Camacho, "Social networks data analysis with semantics: application to the radicalization problem," *J. Ambient Intell. Humaniz. Comput.*, vol. 15, no. 2, pp. 1763–1777, 2024.

[89] S. Malik and D. Gupta, *Assessing the Usage of Various Data Mining Techniques for Analysis of Online Social Networks. InAI-Based Data Analytics*. Auerbach Publications, 2023.

[90] A. S. Rodríguez and E. C. Murillo, "Automatic parameterization of Support Vector Machines for Short Texts Polarity Detection," *CLEI Electronic Journal*, vol. 20, no. 1, pp. 6–7, 2017.

[91] G. Zhao, X. Qian, and X. Xie, "User-service rating prediction by exploring social users' rating behaviors," *IEEE Trans. Multimedia*, vol. 18, no. 3, pp. 496–506, 2016.

[92] H.-J. Choi and C. H. Park, "Emerging topic detection in twitter stream based on high utility pattern mining," *Expert Syst. Appl.*, vol. 115, pp. 27–36, 2019.

[93] Q. Li, A. Nourbakhsh, S. Shah, and X. Liu, "Real-time novel event detection from social media," in *2017 IEEE 33rd International Conference on Data Engineering (ICDE)*, 2017.

[94] H. Jelodar *et al.*, "Latent Dirichlet allocation (LDA) and topic modeling: models, applications, a survey," *Multimed. Tools Appl.*, vol. 78, no. 11, pp. 15169–15211, 2019.

[95] R. Devika, C. Koushik, and V. Subramaniyaswamy, "An event detection on Twitter using ECLAT (equivalence class transformation) algorithm with TRCM (transaction-based rule change

mining)," *International Journal of Pure and Applied Mathematics*, vol. 119, no. 12e, pp. 13347–13356, 2018.

[96] W. He, F.-K. Wang, and V. Akula, "Managing extracted knowledge from big social media data for business decision making," *J. Knowl. Manag.*, vol. 21, no. 2, pp. 275–294, 2017.

[97] X. Zheng *et al.*, "Big Data for Social Transportation," *IEEE Trans. Intell. Transp. Syst.*, vol. 17, no. 3, pp. 620–630, 2016.

[98] E. D'Andrea, P. Ducange, B. Lazzerini, and F. Marcelloni, "Real-time detection of traffic from twitter stream analysis," *IEEE Trans. Intell. Transp. Syst.*, vol. 16, no. 4, pp. 2269–2283, 2015.

[99] Z. Zhou, C. Gao, C. Xu, Y. Zhang, S. Mumtaz, and J. Rodriguez, "Social big-data-based content dissemination in internet of vehicles," *IEEE Trans. Industr. Inform.*, vol. 14, no. 2, pp. 768–777, 2018.

[100] U. Sivarajah, Z. Irani, S. Gupta, and K. Mahroof, "Role of big data and social media analytics for business-to-business sustainability: A participatory web context," *Ind. Mark. Manag.*, vol. 86, pp. 163–179, 2020.

[101] K. Yakobi, B. Scholtz, and B. vom Berg, "A conceptual model of the challenges of social media big data for Citizen e-Participation: A Systematic Review," *Lecture Notes in Computer Science*, pp. 247–259, 2020.

[102] C. Castillo, *Big crisis data: social media in disasters and time-critical situations*. Cambridge University Press, 2016.

[103] K. Cotter and K. Thorson, "Judging value in a time of information cacophony: Young adults, social media, and the messiness of do-it-yourself expertise," *Int. J. Press Polit.*, vol. 27, no. 3, pp. 629–647, 2022.

[104] J. Yin, A. Lampert, M. Cameron, B. Robinson, and R. Power, "Using social media to enhance emergency situation awareness," *IEEE Intell. Syst.*, vol. 27, no. 6, pp. 52–59, 2012.

[105] J. Kim and M. Hastak, "Social network analysis: Characteristics of online social networks after a disaster," *Int. J. Inf. Manage.*, vol. 38, no. 1, pp. 86–96, 2018.

[106] E. Yoo, W. Rand, M. Eftekhar, and E. Rabinovich, "Evaluating information diffusion speed and its determinants in social media networks during humanitarian crises," *J. Oper. Manage.*, vol. 45, no. 1, pp. 123–133, 2016.

[107] S. Choi and B. Bae, "The real-time monitoring system of social big data for Disaster Management," *Lecture Notes in Electrical Engineering*, pp. 809–815, 2015.

[108] K. Lantz, S. Khan, L. B. Ngo, M. Chowdhury, S. Donaher, and A. Apon, "Potentials of online media and location-based big data for urban transit networks in developing countries," *Transp. Res. Rec.*, vol. 2537, no. 1, pp. 52–61, 2015.

[109] J. L. Jimenez-Marquez, I. Gonzalez-Carrasco, J. L. Lopez-Cuadrado, and B. Ruiz-Mezcua, "Towards a big data framework for analyzing social media content," *Int. J. Inf. Manage.*, vol. 44, pp. 1–12, 2019.

[110] S. Manca, L. Caviglione, and J. Raffaghelli, "Big data for social media learning analytics: potentials and challenges," *Journal of e-Learning and Knowledge Society*, vol. 12, 2016.

[111] K. Yamada, H. Takayasu, and M. Takayasu, "Estimation of economic indicator announced by government from social big data," *Entropy (Basel)*, vol. 20, no. 11, p. 852, 2018.

[112] M. Spagnuolo, F. Maggi, and S. Zanero, "BitIodine: Extracting intelligence from the bitcoin network," in *Financial Cryptography and Data Security*, Berlin, Heidelberg: Springer Berlin Heidelberg, 2014, pp. 457–468.

[113] A. K. Sangaiah, A. Goli, E. B. Tirkolaee, M. Ranjbar-Bourani, H. M. Pandey, and W. Zhang, "Big data-driven cognitive computing system for optimization of social media analytics," *IEEE Access*, vol. 8, pp. 82215–82226, 2020.

[114] M. Gupta and J. F. George, "Toward the development of a big data analytics capability," *Inf. Manag.*, vol. 53, no. 8, pp. 1049–1064, 2016.

[115] U. Sivarajah, M. M. Kamal, Z. Irani, and V. Weerakkody, "Critical analysis of Big Data challenges and analytical methods," *J. Bus. Res.*, vol. 70, pp. 263–286, 2017.

[116] J. Fan, F. Han, and H. Liu, "Challenges of Big Data analysis," *Natl. Sci. Rev.*, vol. 1, no. 2, pp. 293–314, 2014.

[117] J. Oliverio, "A survey of social media, Big Data, Data Mining, and analytics," *J. Ind. Integr. Manag.*, vol. 03, no. 03, p. 1850003, 2018.

[118] M.-H. Tsou, "Research challenges and opportunities in mapping social media and Big Data," *Cartogr. Geogr. Inf. Sci.*, vol. 42, no. sup1, pp. 70–74, 2015.

[119] Ishwarappa and J. Anuradha, "A brief introduction on big data 5Vs characteristics and Hadoop technology," *Procedia Comput. Sci.*, vol. 48, pp. 319–324, 2015.

[120] S. Mishra, V. Dhote, G. S. Prajapati, and J. P. Shukla, "Challenges in Big Data Application: A Review," *Int. J. Comput. Appl.*, vol. 121, no. 19, pp. 42–46, 2015.

[121] K. K. Pandey and D. Shukla, "Challenges of big data to big data mining with their processing framework," in *2018 8th International Conference on Communication Systems and Network Technologies (CSNT)*, 2018.

[122] A. B. Alnafoosi and O. Adelakun, "Big Data Adoption Factors and development methodologies: A multiple case study analysis," *Transactions on Computational Science and Computational Intelligence*, pp. 205–223, 2023.

[123] M. Lakoju and A. Serrano, "A strategic approach for visualizing the value of Big Data (SAVV-BIGD) framework," *2016 IEEE International Conference on Big Data (Big Data)*, pp. 1334–1339, 2016.

[124] S. Ren, Y. Zhang, Y. Liu, T. Sakao, D. Huisingh, and C. M. Almeida, "A comprehensive review of big data analytics throughout product lifecycle to support sustainable smart manufacturing: A framework, challenges, and future research directions," *Journal of cleaner production*, vol. 210, pp. 1343–1365, 2019.

[125] A. Yıldız and V. Büyük Verinin Vleri Ve, "Büyük Veri'nin V'leri ve Veri Analitiği," *Pamukkale Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, vol. 2022, no. 51, pp. 361–378.

[126] I. Lee, "Big data: Dimensions, evolution, impacts, and challenges," *Bus. Horiz.*, vol. 60, no. 3, pp. 293–303, 2017.

[127] K. Rahul and R. K. Banyal, "Data life cycle management in big data analytics," *Procedia Comput. Sci.*, vol. 173, pp. 364–371, 2020.

[128] S. Talapatra, N. Ahmed, S. Chakraborty, S. Roy, A. Basu, and A. Guha Sinha, "Problems in data analytics and its solutions," *Journal of Mathematical Sciences & Computational Mathematics*, vol. 1, no. 3, pp. 344–353, 2020.

[129] A. Ganesh, M. Sandhya, and S. Shankar, "A study on fault tolerance methods in Cloud Computing," in *2014 IEEE International Advance Computing Conference (IACC)*, 2014.

[130] M. Risius and R. Beck, "Effectiveness of corporate social media activities in increasing relational outcomes," *Inf. Manag.*, vol. 52, no. 7, pp. 824–839, 2015.

[131] S. Stieglitz, M. Mirbabaie, B. Ross, and C. Neuberger, "Social media analytics-Challenges in topic discovery, data collection, and data preparation. International journal of information management," vol. 39, pp. 156–168, 2018.

[132] A. Specht *et al.*, "Data management challenges in analysis and synthesis in the ecosystem sciences," *Sci. Total Environ.*, vol. 534, pp. 144–158, 2015.

[133] S. Bagga and A. Sharma, "Big data and its challenges: A review," in *2018 4th International Conference on Computing Sciences (ICCS)*, 2018.

[134] E. G. Ularu, F. C. Puican, A. Apostu, and M. Velicanu, "Perspectives on big data and big data analytics," *Database Systems Journal*, vol. 3, no. 4, pp. 3–14, 2012.

[135] S. H. Saleh, R. Ismail, Z. Ibrahim, and N. Hussin, "Issues, challenges and solutions of big data in information management: An overview," *Int. J. Acad. Res. Bus. Soc. Sci.*, vol. 8, no. 12, 2019.

[136] J. Moreno, M. Serrano, and E. Fernández-Medina, "Main issues in Big Data security," *Future Internet*, vol. 8, no. 3, p. 44, 2016.

[137] J. Moura and C. Serrão, *Security and privacy issues of big data. Handbook of research on trends and future directions in big data and Web Intelligence*. IGI Global, 2015.

[138] M. Matali, "An investigation of the impact of social media as an effective communication tool in Namibia: A case study of the affirmative repositioning movement (ARM)," UNAM Repository, https://repository.unam.edu.na/items/59d47e1d-648d-4721-8409-afadf8594440 (accessed Dec. 31, 2024).

[139] B. Mounia and C. Habiba, "Big data privacy in healthcare Moroccan context," *Procedia Computer Science*, vol. 63, pp. 575–580, 2015.

[140] K. U. Jaseena and J. M. David, "Issues, challenges, and solutions: big data mining. CS & IT-CSCP," vol. 4, pp. 131–140, 2014.

[141] S. Kanchi, S. Sandilya, S. Ramkrishna, S. Manjrekar, and A. Vhadgar, "Challenges and solutions in Big Data Management -- an overview," *2015 3rd International Conference on Future Internet of Things and Cloud*, pp. 418–426, Aug. 2015.

[142] A. Sari and M. Karay, "Reactive data security approach and review of data security techniques in wireless networks," *Int. J. Commun. Netw. Syst. Sci.*, vol. 08, no. 13, pp. 567–577, 2015.

[143] B. L. Putro, K. Surendro, and Herbert, "Leadership and culture of data governance for the achievement of higher education goals (Case study: Indonesia University of Education)," 2016.

[144] T. Koltay, "Data governance, data literacy and the management of data quality," *IFLA J.*, vol. 42, no. 4, pp. 303–312, 2016.

[145] E. Curry, "The Big Data Value Chain: Definitions, concepts, and theoretical approaches," *New Horizons for a Data-Driven Economy*, pp. 29–37, 2016.

[146] T. Jukić and M. Merlak, "The use of social networking sites in public administration: The case of Slovenia," *Electronic Journal of E-Government*, vol. 15, pp. 2–18, 2017.

[147] K. Batko and A. Ślęzak, "The use of Big Data Analytics in healthcare," *J. Big Data*, vol. 9, no. 1, p. 3, 2022.

[148] N. Khan *et al.*, "Big data: survey, technologies, opportunities, and challenges," *ScientificWorldJournal*, vol. 2014, p. 712826, 2014.

[149] D. Jayaram, A. K. Manrai, and L. A. Manrai, "Effective use of marketing technology in Eastern Europe: Web analytics, social media, customer analytics, digital campaigns, and mobile applications," *Journal of economics, finance, and administrative science*, vol. 20, pp. 118–132, 2015.

[150] L. S. Matlala, "Improving citizen-based monitoring in South Africa: A social media model," *Afr. Eval. J.*, vol. 12, no. 1, 2024.

[151] K. Kambatla, G. Kollias, V. Kumar, and A. Grama, "Trends in big data analytics," *J. Parallel Distrib. Comput.*, vol. 74, no. 7, pp. 2561–2573, 2014.

[152] S. Vijayaragavan, A. Anand, S. Vignesh, and R. A. X. Annie, "Visualization of big data analysis on social media," in *2017 International Conference on Energy, Communication, Data Analytics and Soft Computing (ICECDS)*, 2017.

[153] A. Ali, J. Qadir, R. ur Rasool, A. Sathiaseelan, and A. Zwitter, "Big Data for Development: Applications and Techniques," *arXiv [cs.CY]*, 2016.

[154] F. Z. Rozony, M. N. Aktar, M. Ashrafuzzaman, and A. Islam, "A Systematic Review of Big Data Integration Challenges and Solutions for Heterogeneous Data Sources," *Academic Journal on Business Administration, Innovation & Sustainability*, vol. 4, no. 04, pp. 1–8, 2024.

[155] W. Xie and K. Karan, "Consumers' privacy concern and privacy protection on social network sites in the era of big data: Empirical evidence from college students," *J. Interact. Advert.*, vol. 19, no. 3, pp. 187–201, 2019.

[156] B. B. Mehta and U. P. Rao, "Privacy-preserving unstructured big data analytics: Issues and challenges," *Procedia Computer Science*, vol. 78, pp. 120–124, 2016.

[157] Z.-G. Liu, X.-Y. Li, and X.-H. Zhu, "scenario modeling for government big data governance decision-making: Chinese experience with public safety services," *Inf. Manag.*, vol. 59, no. 3, p. 103622, 2022.

[158] F. Haneem, N. Kama, N. Taskin, D. Pauleen, and N. A. Abu Bakar, "Determinants of master data management adoption by local government organizations: An empirical study," *Int. J. Inf. Manage.*, vol. 45, pp. 25–43, 2019.

[159] A. Siddiqa *et al.*, "A survey of big data management: Taxonomy and state-of-the-art," b*J. Netw. Comput. Appl.*, vol. 71, pp. 151–166, 2016.

[160] S. Kumar and N. Shah, "False information on web and social media: A survey," *arXiv [cs.SI]*, 2018.

[161] Z. Baraka, *Opportunities to manage big data efficiently and effectively (Doctoral dissertation)*. Dublin Business School, 2014.

[162] K. Yakobi, *A Social Media Analytics Framework for Decision-making in Citizen Relationship Management*. 2022.