

Systematic Review

# A Scoping Review of the Relationship of Big Data Analytics with Context-Based Fake News Detection on Digital Media in Data Age

Khurram Shahzad <sup>1,\*</sup>, Shakeel Ahmad Khan <sup>2</sup>, Shakil Ahmad <sup>3</sup>  and Abid Iqbal <sup>3</sup> 

<sup>1</sup> Department of Library, Government College University, Lahore 54000, Pakistan

<sup>2</sup> Department of Information Management, Islamia University of Bahawalpur, Bahawalpur 63100, Pakistan

<sup>3</sup> Central Library, Prince Sultan University Riyadh, Riyadh 11586, Saudi Arabia

\* Correspondence: khurram@gcu.edu.pk

**Abstract:** The objectives of the study were to identify the relationship between big data analytics with context-based news detection on digital media in the data age, to find out the trending approaches to detect fake news on digital media, and to explore the challenges for constructing quality big data to detect misinformation on social media. Scoping review methodology was applied to carry out a content analysis of 42 peer-reviewed research papers published in 10 world-leading digital databases. Findings revealed a strong positive correlation between quality big data analytics and fake news detection on digital media. Additionally, it was found that artificial intelligence, fact-checking sites, neural networks, and new media literacy are trending techniques to identify correct information in the age of misinformation. Moreover, results manifested that hidden agenda, the volume of fake information on digital media, massive unstructured data, the fast spread of fake news on digital media, and fake user accounts are prevalent challenges to construct authentic big data for detecting false online information on digital media platforms. Theoretically, the study has added valuable literature to the existing body of knowledge by exploring the relationship between big data analytics and context-based fake news on digital media in the data age. This intellectual piece also contributes socially by offering practical recommendations to control the cancer of fake news in society for stopping horrific perils; hence, it has a societal impact. Current research has practical applications for generators of digital media applications, policy-makers, decision-takers, government representatives, civil societies, higher education bodies, media workforce, educationists, and all other stakeholders. Recommendations offered in the paper are a roadmap for framing impactful policies to stay away from the harms of fake digital news.

**Keywords:** big data analytics; context-based fake digital news; digital media age; relation of big data with fake news detection; trending approaches to identify fake news; challenges to build quality big data



**Citation:** Shahzad, K.; Khan, S.A.; Ahmad, S.; Iqbal, A. A Scoping Review of the Relationship of Big Data Analytics with Context-Based Fake News Detection on Digital Media in Data Age. *Sustainability* **2022**, *14*, 14365. <https://doi.org/10.3390/su142114365>

Academic Editor: Andreas Kanavos

Received: 15 October 2022

Accepted: 31 October 2022

Published: 2 November 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Fake news is considered as false reporting of the news originated by self-centered users of social media to mislead the readers intentionally for meeting utilitarian objectives [1]. In the current data age, fake news on digital media is the most prominent social issue that is causing severe dangers and irreparable loss in all fields of life. False information on social media may not easily be identified because bogus facts and figures are intentionally posted to alter public opinions towards certain matters of social significance [2–4]. In the modern times of misinformation, digital social networking websites have proliferated online fake news and the inability to find out accurate information [5,6]. Fake digital content is in full swing due to the emergence of social networking applications, filter bulbs, digitization of human life, machine learning, and deep learning algorithms [7,8]. Due to the quick spread

of online data, fake news flourishes and reaches every corner of the world; consequently, it becomes very difficult to identify correct information from web-based media [9].

Virtual data excess and trending analytics techniques have given birth to big data that “refers to our newfound ability to crunch a vast quantity of information, analyze it instantly, and draw sometimes astonishing conclusions from it [10].” Processed large datasets by leveraging big data are productive in the war against fake news in the age dominated by social networking platforms [11]. Well-established companies of the world digitize products and services to generate big data for knowing their customers’ needs in order to make the right decisions [12,13]. In big data analytics, text mining is a pertinent tool to organize heterogeneous unprocessed data and to extract the correct information from user-generated fake content at new media sites [14]. Quality big data are helpful in detecting fake news on digital media platforms and to stop the spread of false online stories that are disseminated by negative users [15,16]. Big data analytics are a trending practice in the battle against fake news and to identify meaningful information [17,18]. Big data are useful to detect fake information [19] because traditional methods to identify correct information are not sufficient due to the volume and speed of false news in digital media [20]. Big data analytics assist in finding out correct information speedily from large stored data and reducing the harms of fake news being circulated on digital media [21]. Social media big data analytics provides a solution to build an intelligent system to take effective decisions based on correct information [22].

Deep learning architectures are an effective antidote against the fatal disease of online fake news [23]. In the big data age, machine learning algorithms are used to evaluate authenticity of news from large datasets [24]. Construction of propagation patterns prove useful in automatically detecting fake news [25]. Deep learning approaches assist in knowing social media users’ attitudes and identify fake news effectively [26]. Classification of the news based upon artificial intelligence (AI) powered tools is of paramount worth in revealing authenticity of online news [27]. Textual review helps in revealing credibility of the news posted at digital media platforms [28]. Neural networks are of great value to rescue people from disasters of fake information posted at social media applications excessively [29]. Natural language processing technique is a trending method to detect context-based fake news prevalent on social networking websites [30]. Knowledgeable prompt learning is a great tool against fake news posted on the digital media applications to promote baseless and irrational propaganda for personal benefits [31]. Machine learning techniques and quality big data are beneficial to trace the roots of fake news on social media forums [32].

Certain challenges are encountered to identify fake news on digital media including the unavailability of accurate datasets, traditional approaches, and lack of verification attitude [33]. The heterogeneity of a substantial amount of data due to the uncontrollable diffusion of digital media networks causes problems to search for accurate information [34]. During natural calamities and national disaster situations, a huge amount of fake data is dispersed in digital media to create panic among citizens [35]. A single solution does not exist to detect fake information due to its dynamics [36–38]. Detection of fake digital news at an early stage is a worth-mentioning challenge in today’s world of social networks due to the unavailability of processed data [39].

Big data analytics is a phenomenal tool to detect context-based fake news on digital media in the current age dominated by social media platforms through automatic high-tech methods and artificial-based approaches. Instant study aims to find out the relationship between big data analytics with context-based fake news detection on digital media in the data age. In the modern times of fake information posted on digital media forums, the identification of correct news has become a pertinent challenge. This study reveals trending approaches to detect fake news on digital media and manifests practical measures for constructing quality big data to confirm the authenticity of user-posted content in social networking applications. Extant literature illustrated that various studies were carried out on big data and fake news; notwithstanding, a comprehensive scoping review covering

diverse researches conducted in different parts of the world has not been investigated. A scoping review on the relationship between big data analytics and contextual fake news identification based upon substantial empirical investigations held in geographically dispersed lands needs to be carried out. Trending practices displayed via this study will provide new horizons to detect fake news posted on digital media effectively and efficiently. The research also displays challenges being encountered in constructing quality big data to detect misinformation on social media. The study adds significant knowledge to the current body of the literature through a comprehensive scoping review consisting of 42 peer-reviewed research papers. The study also offers social and practical contributions for the decision-takers and policy constructors through the provision of practical solutions to detect fake information on digital media.

### *Research Questions*

The following research questions were addressed in the study:

RQ1. What is the relationship of big data analytics with context-based fake news detection on digital media in the data age?

RQ2. What are the trending approaches to detect fake news on digital media?

RQ3. Which are the challenges for constructing quality big data to detect misinformation on social media?

## **2. Methodology**

The researchers applied the “Preferred Reporting Items for the Systematic Review and Meta-analysis” (PRISMA) procedures to conduct the study. “PRISMA is an evidence-based minimum set of items for reporting in systematic review and meta-analysis. PRISMA is used for reporting of review, evaluating randomized trials, but it can also be used as a basis for reporting systematic review” [40]. Having applied this methodology, Shahzad and Khan [41] conducted a systematic review of the factors leading to the implementation of semantic digital libraries. PRISMA is based upon four main parts along many steps at each part. The first part is planning, which covers focused research questions and search strategy. The second part is the selection which is aimed to extrapolate and sort the data. The third part is extraction that is carried out for evaluating the data through a pre-set systematic assessment. The last stage known as data synthesis is applied to analyze the data for producing successive procedures. These four parts are applied in this study and elaborated below:

### **A. Phase 1: Planning**

#### **(1) Focused research questions**

The focused research questions of the current study include the relationship of big data analytics with context-based fake news detection on digital media in the data age, trending approaches to detect fake news on digital media, and the challenges for constructing quality big data to detect misinformation on social media.

#### **(2) Search strategy**

Strategies used to search required terms, sources to find and locate literature, and the procedure of the search have been detailed below:

##### **a: Search terms**

Search terms of the study were created via pre-set methods and criteria. The following ways were adopted to retrieve the most matching literature at par with the set-focused research questions:

Use of key variables from article-title as a major technique during the finding of required content.

Shaping a general research question of the study.

Selection of some constructs from the pre-developed study questions showing clear directions.

Follow keywords applied by other authors in their papers.

Creation of synonyms-list to explore the literature further.

Employment of Boolean operators “OR”, “AND”, and “NOT” to retrieve refined, and precise results.

The search was held through diverse techniques to access the maximum number of relevant documents. The following search phrases were used for exploring the most matching results keeping in view focused research questions:

(“Big data” OR “Big data analytics” OR “Relation of big data with fake news detection” OR “Methods to detect fake contextual news” OR “Challenges to detect fake news”) OR “Role of data age in the spread of fake online news” OR “Impact of big data on fake news identification” OR “Challenges to create big data” OR “Effects of digital sites in fake news diffusion” OR “Big data analytics” OR “Digital media” OR “Fake news on social media” OR “Context based fake news detection” OR “Fake news detection tool” OR “Problems to generate quality big data” OR “Machine learning” AND “Fake news detection” AND “Social media” AND “Big data” AND “Big data” AND “Digital fake news control” AND “Data age” AND “Fake online content” AND “Robust fake news detection techniques” AND “Big data analytics” AND “Social networks” AND “Fake news in networked age” AND “Fake news data analysis” AND “Quality data for fake news control” AND “Fake news detection on social media” AND “Fake news detection problems” AND “Big data approaches” AND “Control of false online news” AND “Big data” AND “Modern journalism” OR “Data age” OR “Data journalism” AND “Technological approaches” OR “Identification of fake news” AND “Big data framework” OR “Analysis of social media content” AND “Solutions to combat fake digital news” AND “Harvesting big data” OR “Combating user-generated online content” AND “Fake news on social media” AND “Deep learning” AND “Contextual fake news detection” AND “Big data analytics” AND “Social media platforms” AND “Social context”) (“Fake news on social media” NOT “Traditional media”, “Relation of big data with fake news detection” NOT “Printing press”, “Big data analytics” NOT “Traditional communication media”

b: Use of literature resources and existing research

The authors used the world’s 10 leading digital databases to conduct an in-depth search: Web of Science, Scopus, Emerald, Summon, Elsevier, Google Scholar, Taylor & Francis, Pro-Quest, Wiley Inter-Science, and IEEE Xplore. Restrictive phrasing was used for accessing the required results in accordance with the pre-formulated research questions. Advance search options were utilized to retrieve the most relevant and narrow results. Articles published in peer-review impact score journals ranging from 2015 to 2022 were included to conduct the scoping review.

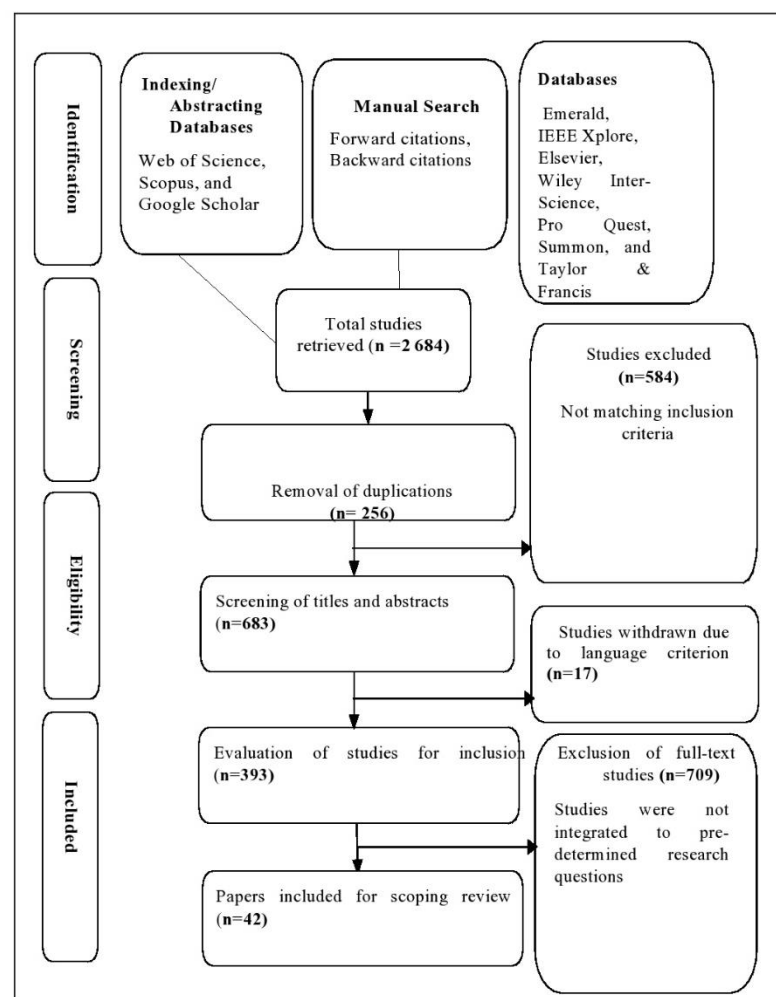
## B. Phase 2: Selection

### (1) Search process

A comprehensive search was done to find and locate all existing relevant literature. Figure 1 provides a graphical description of multiple steps that were applied in that procedure.

Step 1: Ten renowned electronic databases were considered to retrieve the desired documents.

Step 2: For avoiding duplications, scrutiny was held of all the existing content. Non-matching manuscripts were excluded from the list. To ensure relevancy, articles’ titles were observed carefully. Outdated articles were not added to the study. A total of 2684 documents were found while 455 articles were shortlisted after the removal of duplications and irrelevant results. Through the screening process, 974 articles were removed. The authors applied pre-developed criteria to choose papers aligned with focused research questions. Resultantly, 42 papers were selected due to alignment and integration with the focused study research questions.



**Figure 1.** Diagram of the search process.

## (2) Scrutiny and filtering

For ensuring relevancy, 2684 retrieved documents were filtered and analyzed. Multiple techniques were carried out to execute the process. A critical analysis of the papers' titles was undertaken to conduct scoping review of the latest relevant documents. The language of the selected articles was English. Only research papers were selected to conduct scoping review while other types of publications were not added to the current paper. Recently published papers were preferred while outdated manuscripts were not included in the list.

### C. Phase 3: Extraction

A score was given to the accessed articles. The score was provided keeping in view the most closely related research questions. Studies meeting the set criteria were provided a score. The procedure enabled authors to withdraw 2642 documents and to include 42 of the most relevant and the most focused research papers.

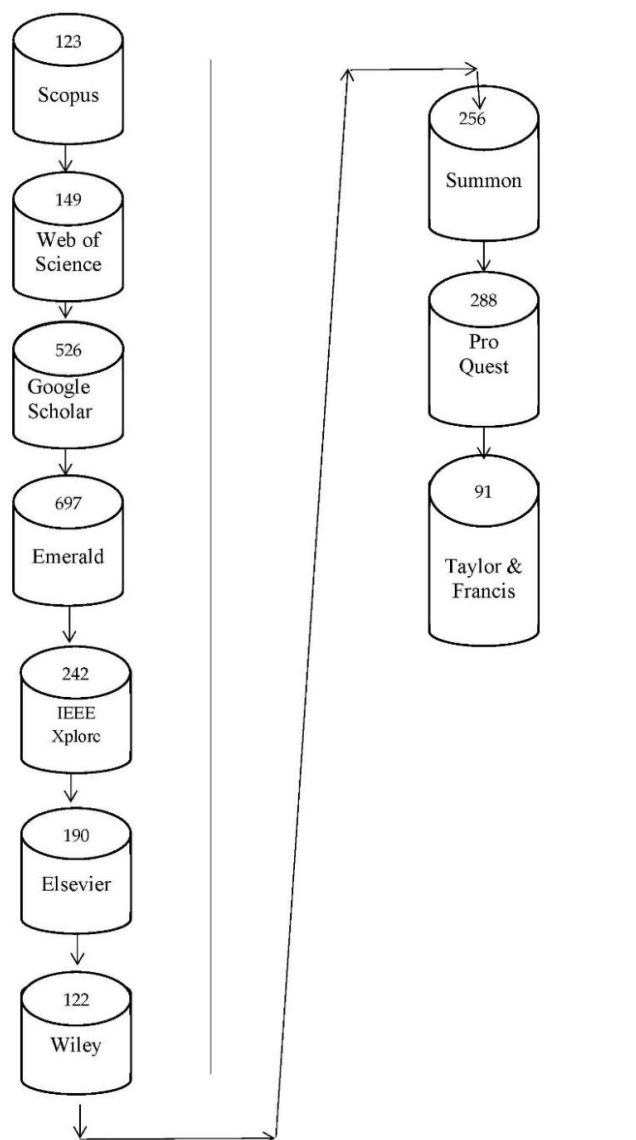
### D. Phase 4: Execution

The validity of the articles was checked to ensure validity through strict evaluation of the list against pre-determined eligibility criteria. Papers published before 2015 were excluded from the list. Most relevant papers were added to the study via critical evaluation. Papers having no similarity with the study research questions were excluded.

### 3. Results

#### 3.1. An Overview of the Selected Studies

On the whole, 2684 manuscripts were accessed through the world's ten leading digital databases and tools: Scopus (123), Web of Science (149), Google Scholar (526), Emerald (697), IEEE Xplore (242), Elsevier (190), Wiley Inter Science (122), Summon (256), Pro-Quest (288), and Taylor & Francis (91). These documents were downloaded from March 2022 to June 2022. In total, 42 research papers published in peer-reviewed journals were chosen to carry out the current study. Figure 2 displays the breakdown of the accessed publications from the above-cited ten digital databases and tools:

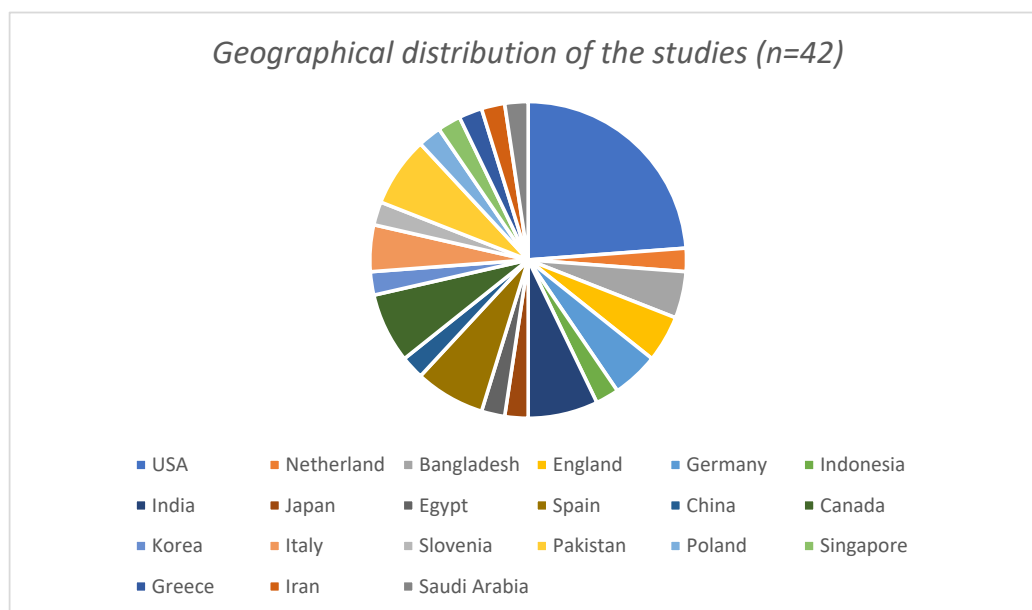


**Figure 2.** The search process.

#### 3.2. Geographical Distribution of the Studies

Figure 3 manifests the geographical territories of the research papers selected to carry out systematic review. Results revealed that studies had been investigated in 21 different regions across the world. It was found that the United States of America was on the top with 10 documents, whereas Pakistan, Canada, and India were in the second spot regarding research output in the area of big data and fake news, while England, Germany, Bangladesh, and Italy were in the third slot. It is important to mention that other countries ( $n = 12$ ) had produced one article each. Selected papers represent a vast range of geographically

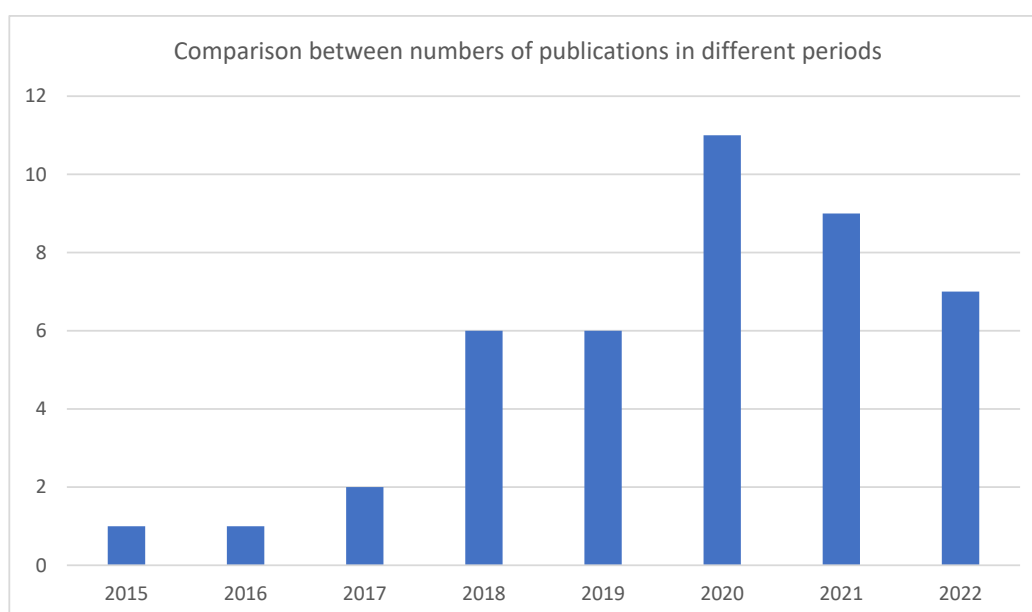
dispersed localities. It is also worth mentioning that all selected documents ( $n = 42$ ) had been published in different journals.



**Figure 3.** Geographical distribution of the studies ( $n = 42$ ).

### 3.3. Years Trends of the Selected Studies

A comparison analysis was conducted to show comparison between numbers of publications in the periods from 2015 to 2018 with the period from 2019 to 2022. It was found that during 2015 to 2018, only 10 studies had been conducted related to the research topic. Nonetheless, 32 papers had been produced related to the big data and fake news during 2019 to 2022. It shows that in the recent years, big data analytics and fake news are emerging areas for the investigators. Figure 4 reveals graphical depiction of the comparison between numbers of publications in the periods from 2015 to 2018 with the period from 2019 to 2022.

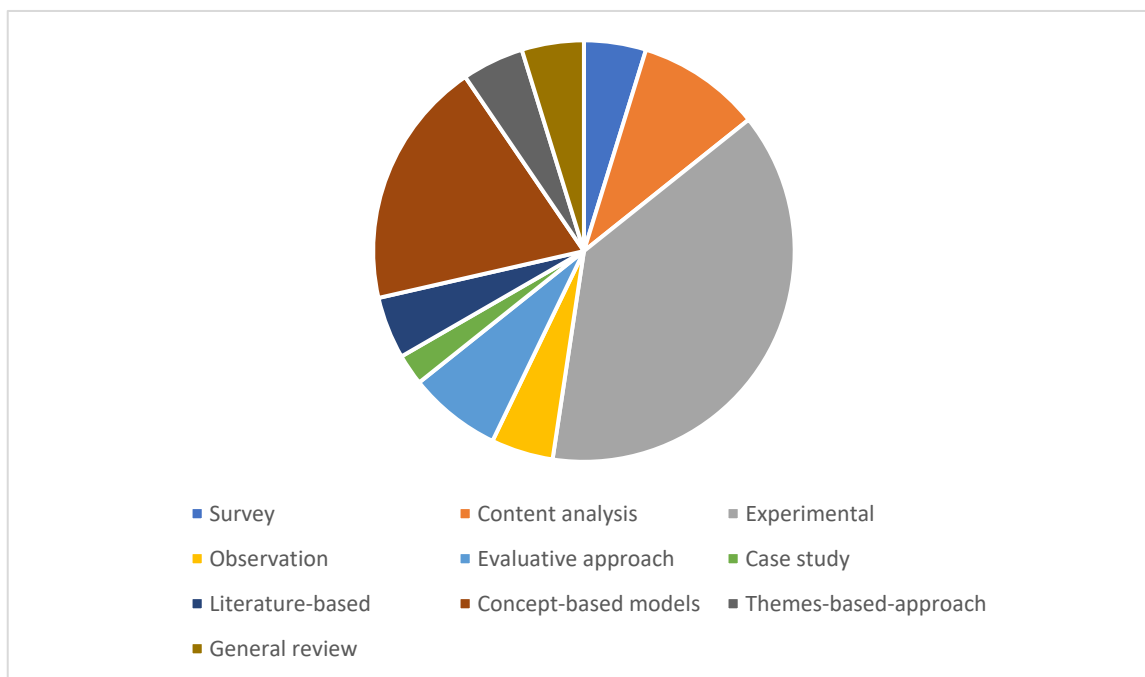


**Figure 4.** Comparison between numbers of publications in the periods from 2015 to 2018 with the period from 2019 to 2022.



### 3.4. Research Methodologies of the Previous Studies

Figure 5 shows descriptive analysis of different research methodologies that were applied in the selected manuscripts ( $n = 42$ ). Analysis revealed that the majority of the investigators working in the area of big data and fake news had used experimental research method ( $n = 16$ ). The second most applied methodology included concept-based models ( $n = 8$ ), while the third top used method was of content analysis ( $n = 4$ ). Findings of the study showed that 10 different research methodologies had been applied by the researchers in 42 different studies.



**Figure 5.** Research methodologies of the previous studies.

The findings of the study, based on focused research questions, are detailed as follows.

### 3.5. Relationship between Big Data Analytics with Context-Based Fake News Detection

In 19 studies out of 42, a positive relationship was identified between big data analytics and context-based fake news detection on digital media in the data age (Table A1). Different authors concluded through empirical investigations that big data analytics were an antidote against the fatal disease of fake news spreading rapidly on social media. Lewis and Westlund [42] proved that big data analytics and fake news detection were positively correlated with each other. Bates et al. [15] identified that big data improved accuracy in health-related information; consequently, correct information was used to make certain decisions. Olmedilla et al. [12] remarked that big data was of paramount worth in detecting accurate information from online user-generated content. Guo and Vargo [43] mentioned that the correlation between big data analytics and fake news detection was positively significant.

Golbeck et al. [2] maintained that the big dataset was useful to the research community and in understanding the nature of fake news and ways of fighting it. Torabi and Taboada [16] reflected that large data sets confirmed news credibility and saved from the social harms of fake news cancer. Mahabub [9] displayed that authentic big data was positively associated with fake news detection in the networked world. Nakamura et al. [33] claimed that big data analytics could be used to advance efforts to combat the ever-growing, rampant spread of disinformation in today's society. Khan et al. [6] asserted that big data detected fake information on social networking sites in the current data age. Hassani et al. [14]



illustrated that text mining in big data analytics was a powerful tool against fake news on digital media. Ianni et al. [34] highlighted that big data analytics helped in analyzing social network data to retrieve correct information. Jung et al. [8] discovered that big data analytics uncovered digital fake news and led toward existing ground realities. Kauffmann et al. [17] observed that big data led to contextual fake news detection on social networking applications.

King and Wang [18] noted that a big data-driven approach found the validity of online posted news. Supriyanto et al. [21] argued that big data assisted in using correct and fast data from anywhere safely and conveniently. Murayama [44] concluded that a big dataset assessed the truthfulness of a certain piece of news from news content being posted at digital media forums. Darwiesh et al. [22] inferred that social media big data analytics was a promised solution to develop classical business intelligence systems for detecting false online news. Raza and Ding [39] recommended that big data sets proved valuable in fake news identification in modern times of technological innovations. Chauhan and Palivela [29] indicated that an ensemble-based deep learning model classified online information as real or fake for an easy identification of fake news from large datasets.

### 3.6. Trending Approaches to Detect Fake News on Digital Media

In light of evidence-based data (Table A1), five trending approaches were discovered to detect fake news on digital media. The approaches included artificial intelligence, fact-checking sites, neural networks, new media literacy, and miscellaneous trends. These trending approaches are interpreted as below alternately:

#### 3.7. Artificial Intelligence

Using automatic machine learning classification models is an efficient way to combat the widespread dissemination of fake news [33]. Ensemble voting classifier based; an intelligent detection system is used to deal with news classification for both real and fake tasks. Machine-learning algorithms like naive bayes, K-NN, SVM, random forest, artificial networks, logistic regression, gradient boosting, and Ada boosting, etc. are used for fake news detection [9]. Artificial intelligence, natural language processing, and machine learning approaches are effective to identify fake online news [27]. Generative machine learning, artificial networks, and artificial intelligence tools are trending means to detect fake information on digital media [7,36,37].

#### 3.8. Fact-Checking Sites

Fact-checking is a trending approach to combat with fake information on digital media platforms [2,33,45]. Fact-checking websites examine the news source to check the authenticity and accuracy of the online news [16]. Real-life fact-checking websites and fact verification datasets offered practical solutions to display the originality of the web-based news [19,44]. Automatic fake news detectors were highly instrumental in the war against digital fake news [17]. Fact-checking systems, and an automatic fake news detection approach in chrome environment through contingent evaluation methods, provided evidence-based facts [26,35].

#### 3.9. Neural Networks

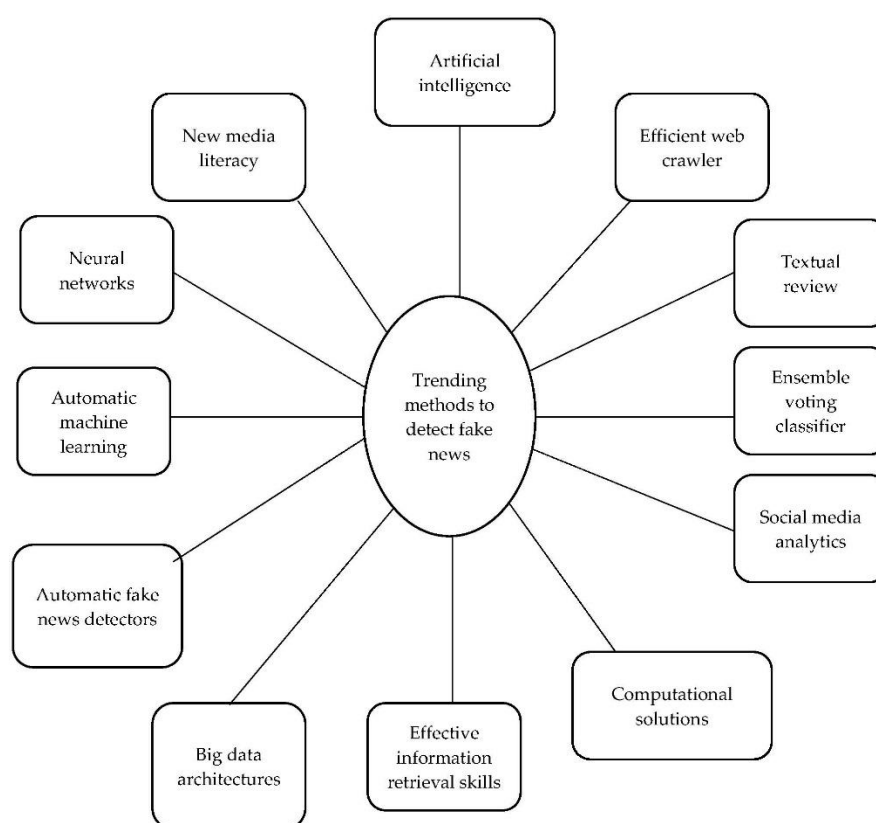
Deep learning models and architectures, neural networks, and natural language processing facilitate in detecting fake news for stopping pernicious news on digital media [6,16,23]. Classification-based models, blockchain-based frameworks, machine learning, big data architectures, machine learning ensemble approach, and natural language processing technology are trending techniques for fake news prevention [5,7,13,19,24,29,38]. Machine learning, deep learning methods, and real-world datasets are a productive source to find out fake news from the flood of misinformation [31,32].

### 3.10. New Media Literacy

New media literacy is a pertinent technique to control fake news perils on digital media platforms [34,44]. The usage of official sources leads to the deletion of rumor-related content [8]. Effective information retrieval skills are fruitful in finding out accurate information [38]. Textual review, data classification, and text analysis are useful in revealing false information from digital media platforms [13,37].

### 3.11. Miscellaneous Trends

Some other pertinent trending techniques to detect contextual fake news on digital media include image features supply models, social media analytics, IQ-based tools, personality traits [20,33,34], and digital media content analysis, effective web-crawlers, computational solutions, identification of users' profiles, and sentiments analysis tools [12,13,17,26,39,46]. Figure 6 displays a graphical depiction of trending approaches being applied to detect fake news on digital media.



**Figure 6.** Trending approaches to detect fake news.

### 3.12. Challenges for Constructing Quality Big Data to Detect Misinformation on Social Media

The study manifested five major challenges that were encountered while constructing quality big data to detect misinformation on social media (Table A1). Challenges were hidden agendas, the volume of fake information on digital media, massive unstructured data, the fast spread of fake news on digital media, and fake user accounts. These main challenges further classified into sub-challenges covering integrated themes are elaborated below:

### 3.13. Hidden Agenda

In social media forums, individuals and institutions have certain hidden agendas and they transmit hidden strategies to attract others for attaining set objectives [43]. The complex nature of fake news and social media comments accompanied by doubtful images or videos create suspicions in viewers' minds [33,44]. Social media data possess misinformation, fake

accounts, and fake news [22]. Fake data spreads faster and penetrates social networks to a larger extent than credible news [16]. Cyberbullying is progressively turning into a typical issue that is causing unmanageable problems these days [9]. Conspiracy and fake sites promote hidden agendas for the interests of certain people and organizations [19]. There is the manipulation of facts via personal emotions, and unchecked user-generated content [36]. The negative role of journalists and YouTubers is reshaping the media landscape and promoting false doctrines in society [4]. Biased opinions and propagation patterns cause an obstacle for the automatic fake news detection [25,27].

### 3.14. Volume of Fake Information on Digital Media

Industry 4.0 is creating more data than ever before in mankind-history [13]. The structure of international information is not balanced as online news is generated on a massive scale [43]. Big data have a large volume and usually consist of both qualitative and quantitative components from a variety of data types [47]. A huge number of posts are generated on social networking systems [44]. An unprecedented amount of heterogeneous data, the large amount of user-generated data, the high-speed generation rate, and excessive usage of popular social networks cause issues in the creation of quality of big data to detect contextual fake news [34]. The vast amount of content on digital media, huge user-generated content, ideological polarization, and decreasing trust in traditional media create problems in quality big data creation [3,8]. Diverse sources of conflicting information put hurdles in the detection of context-based fake news in digital media [5]. A huge amount of contextual data, the volume of data in the global data sphere, and the wide dissemination of fake news on social media applications make it difficult to identify the accuracy of the news [12,20,46]. Huge volumes of fake news posted by malicious users, and diffusion of low-quality news in social media, are serious challenges to detect context-based fake news in the current era of disinformation [28,32].

### 3.15. Massive Unstructured Data

Digital data have a big drawback concerning data quality because they do not cover the whole population [47]. A lack of effective, comprehensive datasets has been a problem for fake news research and detection model development [33]. Massive and unstructured data on social media within a short time-span and building effective gathering data tools are big challenges due to the different structures, types, and the huge amount and velocity of creation data on social media platforms. As data are unstructured and collected from a wide range of users, the quality of data will be decreased [22]. Challenges of content evaluation, changing users' behaviors, overflowing of information resources, unmanageable spammy content, and shortage of labeled data are barriers to identifying online news authenticity [3,20].

### 3.16. Fast Speed of Fake News on Digital Media

In internet-based life, data is spreading quickly [9]. The wide spread of fake news and the speed and extent of the spread of fake information on social media are certain challenges to find out correct news on digital media [9,16]. Fake news spreads on social media and is perhaps more popular than ever [45]. High speed of fake news proliferation on social media, misinformation at digital sites, and Infodemic are pertinent challenges [19,20,35]. There is an easy spread of fake news on social media due to networked affordances, and the digitization of human life via social networking applications are significant drives for the unstoppable proliferation of fake news [3,7]. Digital journalism, sensational news for an increased rating, the fast reach of online content, and the lack of comprehensive and community-driven fake news data sets are obvious problems in confirming the credibility of digital news [4,46]. The rapid adoption of social media platforms is, indeed, a great challenge to identify fake news [24].

### 3.17. Fake User Accounts

Social bots significantly contribute to fake information [36]. Fake profile trends, security issues, and fake user accounts make it difficult to detect fake news at an early stage [3,39,48].

## 4. Discussion and Implications

This study is the first scoping review in the area of contextual fake news detection on digital media via big data analytics. The findings of the research are based on 42 peer-reviewed research papers published in the world's leading digital databases. The selected studies ( $n = 42$ ) were published in the English language and investigated in geographically dispersed regions of the world. Extracted data illustrated that there was a strong positive relationship between big data analytics and contextual fake news detection in digital media in the current data age. Evidence-based data sets also manifested trending tools to identify fake news on social media applications and challenges being encountered in constructing quality big data to detect misinformation on digital media forums.

Big data analytics is a phenomenal weapon in the battle against fake online information that is disseminated by evil-minded social site users for meeting hidden objectives. The instant study revealed that in the modern data age, a positive correlation existed between big data analytics and contextual fake news detection on digital media platforms provided that quality data is generated. Content analysis of the selected studies for scoping review manifested that text mining in big data analytics, big data sets, big quality data, social media big data, large datasets, and authentic big data assist in analyzing content posted at digital media applications and reveal the authenticity of the online information. Without quality big data analytics, contextual fake news on social media may not be traced. Hence, accurate content generation is of paramount worth in capturing fake news on digital media forums. In the modern age led by social media applications, online fake news is a great challenge; therefore, big data analytics are highly significant to identify correct information from the flood of misinformation effectively. Lewis and Westlund [42], Olmedilla et al. [12], Guo and Vargo [43], Golbeck et al. [2], Khan et al. [6], Jung et al. [8], King and Wang [18], and Darwiesh et al. [22] also reported similar results in their studies.

Some pertinent trending approaches are applied to detect fake news on digital media in the current data age. Artificial intelligence (AI) is a significant approach to identify fake news from social media in modern times of misinformation. AI-powered tools assist in stopping the diffusion of fake news on digital media forums and to reveal the truthiness of online posted news. Automatic intelligent detection systems are utilized for contextual fake news identification. This evidence is in line with the findings illustrated by Mahabub [9] and Kozik et al. [37] in their empirical studies. Fact-checking sites are also an effective trend to identify contextual fake news. Real-life fact-checking websites examine the originality of online news through automatic rigorous evaluation methods and modern-driven techniques. This result is at par with the findings of Golbeck et al. [2], Nakamura et al. [33], Murayama [44], and Jo et al. [35]. Neural networks based upon blockchain applications, deep learning, and classification models support in bringing out correct news from the flood of misinformation being disseminated by non-serious social sites users. This outcome is linked with the results displayed by Huckle and White [5], Khan et al. [6], Marquez et al. [13], Meesad [38], and Qayyum et al. [7] in their articles. New media literacy, civic literacy, efficient information retrieval expertise, text analysis, confirmation of digital content from authentic sources, and verification attitude before posting the news on digital media networks guide digital users to differentiate between fake and correct information. This illustration is integrated into the results reported by Marquez et al. [13], Ianni et al. [34], and Jung et al. [8] in their investigations. Other note-worthy trends to find out the accuracy of digital news include social media analytics, effective search engines, efficient retrieval systems, and human emotions analysis tools. Similar trends to detect online fake news were concluded by Olmedilla et al. [12], Kauffmann et al. [17], Shu et al. [46], Nakamura et al. [33], Zrnec et al. [20], and Raza and Ding [39] in their scholarly contributions.

Challenges of a diverse nature are faced to construct quality big data to detect misinformation on social media. The hidden agenda of particular individuals and groups is a big challenge to develop authentic content for the identification of correct news from digital media platforms. Unending suspicious comments, cyberbullying, conspiracy, fake sites, self-centered YouTubers, and journalists cause difficulties in the construction of quality metadata for finding out contextual fake information on social media applications. These reflections confirm the findings displayed by Guo and Vargo [43], Veglis and Maniou [4], and Torabi and Taboada [16] in their works. A huge amount of massy data on social networking websites is a significant obstacle to the creation of quality big data. The heterogeneity of the data due to users' autonomy to post any content on digital media causes difficulties to create quality datasets. The vast amount of user-generated content on social media applications is a prominent cause for the unavailability of authentic big datasets for detecting fake news from the heaps of misinformation. The plurality of thoughts posted on digital media forums is a great hindrance to developing quality metadata to find out correct and original news. Uncontrollable diverse context text in the global data sphere makes it extremely complex to display attested information. These outlooks match with the results of the studies investigated by Olmedilla et al. [12], Huckle and White [5], Al-Rawi et al. [3], Baur et al. [47], Shu et al. [46], and Zrnec et al. [20]. Massive unstructured data is also an obstacle to create quality big data for capturing fake news from diverse sources. The unavailability of authentic datasets is an obvious reason not to build quality big data for stopping the spread of fake-information-flood at digital media sites. This result is in accordance with the result of Darwiesh et al. [22] who mentioned that the lack of labeled data was a problem to construct quality big data. The speedy proliferation of fake news on social media due to technological advancements and the affordability of digital tools lead to the unavailability of quality big datasets to retrieve correct news. This finding is related to the studies conducted by Veglis and Maniou [4], Mahabub [9], Shu et al. [46], and Qayyum et al. [7]. This study also revealed that social bots also contributed a substantial amount of fake information. A similar conclusion was presented by Liu [36], Al-Rawi et al. [3], Awan et al. [48], and Raza and Ding [39] through their studies.

Theoretically, the current study has added valuable literature to the existing body of knowledge by exploring the relationship between big data analytics and context-based fake news on digital media in the data age. This intellectual piece also contributes socially by offering practical recommendations to control the cancer of fake news in society for stopping horrific perils, hence it has a societal impact. Current research has practical applications for generators of digital media applications, policy-makers, decision-takers, government representatives, civil societies, higher education bodies, media workforce, educationists, and all other stakeholders. The study manifests trending approaches to identify correct news and avoid fake information from digital media. It offers practical measures to construct quality big data for bringing out authentic news from credible sources. Recommendations offered in the paper are a roadmap for framing impactful policies to stay away from the harms of fake digital news.

## 5. Conclusions and Recommendations

In light of the content analysis of the 42 studies, it is concluded that a positive relationship exists between big data analytics and context-based fake news detection on digital media. Quality big data analytics assists in identifying fake news on social media applications. The study has displayed five key trends (artificial intelligence, fact-checking sites, neural networks, new media literacy, and miscellaneous approaches) supported by several sub-themes to identify fake news on the digital media platforms and also five major challenges (hidden agenda, volume of fake information on digital media, massive unstructured data, fast spread of fake news on digital media, and fake user accounts) further classified into sub-challenges to construct quality big data for verifying the authenticity of the online news.

The following applicable recommendations are offered in light of evidence-based findings:

- An innovative course on big data, covering diverse dimensions, should be taught in library schools for spreading awareness and necessary skills to identify contextual fake news on digital media platforms. There should be a strong positive liaison between library schools and the industry to develop need-based content for imparting creative learning and to provide skilled workers in the market.
- Digital media generators should take strict measures against all those users who post hidden agendas to prevail over irrational practices to shake foundations of the society.
- Adequate steps should be executed to control heterogeneity, volume, and pace of unstructured data for stopping fake news diffusion on digital media.
- Fake accounts should be banned permanently from digital media sites so the amount of posted content may be minimized.
- Quality big data and social media metadata should be developed for detecting context-based fake news.
- New media literacy skills should be infused in web users so that they may verify the originality of the news before posting on digital media applications.
- Artificial intelligence-powered tools should be applied for automatically detecting fake online news effectively and efficiently.
- Government and higher education bodies should plan and execute all necessary steps for implementing, maintaining, and sustaining quality big digital media content for the immediate detection of context-based fake news on social media applications.

## 6. Limitations and Future Research Directions

The study has certain limitations in spite of significant theoretical, practical, and social contributions. A pertinent limitation of the current study is the inclusion of only articles ( $n = 42$ ) to carry out systematic review for constructing an evidence-based framework to control fake news diffusion on the digital media for constructing impactful policies to control the cancer of fake news on the digital media. Other types of documents, i.e., magazines, books, conference proceedings, dissertations, newsletters, grey literature, government documents etc. have not been included. Another worth-mentioning limitation is the inclusion of only those papers that were published in the English language. Current study has explored the relationship of big data analytics with context-based fake news detection on digital media in data age. Future investigators might conduct the relationship between new media literacy and web-based fake news epidemic control. Researchers of the future should also empirically test the results of our study by considering varying cultural traditions regarding fake news sharing on social media. A future study might also be conducted through scoping review on the relationship between emotions management and fake news resistance on the digital media.

**Author Contributions:** Conceptualization, K.S., S.A.K. and A.I.; methodology, K.S. and S.A.K.; validation, A.I., S.A.; formal analysis, K.S., A.I. and S.A.; investigation, K.S. and A.I.; resources, S.A.; data curation, S.A. and A.I.; writing—original draft preparation, K.S., S.A.K., A.I. and S.A.; writing—review and editing, K.S. and A.I.; visualization, K.S. and A.I.; supervision, S.A.K.; project administration, S.A.K. and A.I.; funding acquisition, S.A. and A.I. All authors have read and agreed to the published version of the manuscript.

**Funding:** This project was financially supported by Prince Sultan University Riyadh, Saudi Arabia for the provision of APC.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Not applicable.

**Acknowledgments:** The authors acknowledge the financial support of Prince Sultan University Riyadh, Saudi Arabia for the provision of APC.

**Conflicts of Interest:** The authors declare no conflict of interest.

## Appendix A

**Table A1.** Data extracted from 42 research articles.

S.N.	Author	Year	Country	Journal	Relation of Big Data Analytics with Fake News Detection	Trending Approaches to Detect Fake News on Digital Media	Challenges for Constructing Quality Big Data to Detect Misinformation on Social Media
1.	Vargo and Amazeen	2018	USA	New media & society		Fact checking	Fake news spreads on social media and is perhaps more popular than ever.
2.	Guo and Vargo	2017	USA	Journal of Communication	Correlation between big data analytics and fake news detection is significant.		<ul style="list-style-type: none"> <li>• The structure of international information is not balanced.</li> <li>• Generation and consumption of online information to a massive scale</li> <li>• Hidden agenda setting on social media forums</li> </ul>
3.	Baur et al.	2020	Germany	Historical Social Research/Historische Sozialforschung			<ul style="list-style-type: none"> <li>• Digital data have a big drawback concerning data quality because they do not cover the whole population</li> <li>• Big data have a large volume.</li> <li>• Big data sets usually consist of both qualitative and quantitative components from a variety of data types (e.g., numerical, verbal, and visual data) and data sources.</li> <li>• There are different types of data within the internet.</li> </ul>
4.	Golbeck et al.	2018	Netherlands	WebSci	Big dataset is useful to the research community and on understanding the nature of fake news and ways of fighting it.	Automated system for fake news detection.	



Table A1. Cont.

S.N.	Author	Year	Country	Journal	Relation of Big Data Analytics with Fake News Detection	Trending Approaches to Detect Fake News on Digital Media	Challenges for Constructing Quality Big Data to Detect Misinformation on Social Media
5.	Nakamura et al.	2020	USA	arXiv preprint arXiv	Big data analytics can be used to advance efforts to combat the ever-growing rampant spread of disinformation in today's society.	<ul style="list-style-type: none"> <li>Using automatic machine learning classification models is an efficient way to combat the widespread dissemination of fake news.</li> <li>Image features supply models with more data that can help immensely to identify fake images and news that have image data.</li> <li>Fact checking</li> </ul>	<ul style="list-style-type: none"> <li>A lack of effective, comprehensive datasets has been a problem for fake news research and detection model development.</li> <li>Diversity of fake news</li> </ul>
6.	Khan et al.	2019	Bangladesh	Machine Learning with Applications	Big data detects fake information.	<ul style="list-style-type: none"> <li>Machine learning approaches</li> <li>Neural networks to detect fake news</li> <li>Deep learning models</li> </ul>	
7.	Supriyanto et al.	2021	Indonesia	Paedagoria: Jurnal Kajian, Penelitian dan Pengembangan Kependidikan	With big data we can use the correct and fast data from anywhere safely and conveniently.		
8.	Murayama	2021	Japan	arXiv preprint arXiv	Big dataset assesses the truthfulness of a certain piece of news from news content	<ul style="list-style-type: none"> <li>Fact checking sites</li> <li>Fact verification datasets</li> <li>News literacy</li> </ul>	<ul style="list-style-type: none"> <li>Diversity and complex nature of fake news</li> <li>Huge amount of posts on social networking systems</li> <li>Veracity estimation for social media comments accompanied by doubtful images or videos.</li> </ul>

Table A1. Cont.

S.N.	Author	Year	Country	Journal	Relation of Big Data Analytics with Fake News Detection	Trending Approaches to Detect Fake News on Digital Media	Challenges for Constructing Quality Big Data to Detect Misinformation on Social Media
9.	Darwiesh et al.	2022	Egypt	Journal of Healthcare Engineering	<ul style="list-style-type: none"> <li>• Social media big data analytics is a promised solution to develop classical business intelligence systems.</li> <li>• Using big data analytics finds out fake information on the digital media.</li> </ul>		<ul style="list-style-type: none"> <li>• Massive and unstructured data on social media within a short time-span</li> <li>• Building effective data gathering tools is a big challenge due to the different structures, types, and the huge amount and the velocity of creation data on social media platforms.</li> <li>• As data are unstructured and collected from a wide range of users, the quality of data will be decreased.</li> <li>• Social media data may face some issues such as misinformation, fake accounts, and fake news. These issues make a bad effect on any analytical process, and the output insights will be biased.</li> </ul>
10.	Torabi and Taboada	2019	Canada	Big Data & Society	Large data sets confirm news credibility.	<ul style="list-style-type: none"> <li>• Natural language processing</li> <li>• To examine the news-source</li> <li>• Automatic fact checking and classification</li> <li>• To educate the people for stopping pernicious news</li> <li>• Fact checking websites</li> </ul>	<ul style="list-style-type: none"> <li>• Speed and extent of the spread of the fake information on social media.</li> <li>• Fake data spreads faster and penetrates social networks to a larger extent than credible news.</li> </ul>

Table A1. Cont.

S.N.	Author	Year	Country	Journal	Relation of Big Data Analytics with Fake News Detection	Trending Approaches to Detect Fake News on Digital Media	Challenges for Constructing Quality Big Data to Detect Misinformation on Social Media
11.	Mahabub	2020	Bangladesh	SN Applied Sciences	Authentic big data is positively associated with fake news detection.	<ul style="list-style-type: none"> <li>Ensemble Voting Classifier based, an intelligent detection system is used to deal with news classification both real and fake tasks.</li> <li>Machine-learning algorithms like Naïve Bayes, K-NN, SVM, Random Forest, Artificial Neural Network, Logistic Regression, Gradient Boosting, Ada Boosting, etc. are used for fake news detection.</li> </ul>	<ul style="list-style-type: none"> <li>Widespread of fake news</li> <li>In the internet-based life, the data are spread quickly and, subsequently, discovery components ought to almost certainly foresee news quickly enough to stop the dispersal of fake news.</li> <li>Cyberbullying is progressively turning into a typical issue among adolescents these days.</li> </ul>
12.	Ianni et al.	2020	Italy	Journal of Intelligent Information Systems	Big data analytics assist in analyzing the social networks data.	<ul style="list-style-type: none"> <li>Social media analytics to analyze the online data</li> <li>New media literacy</li> </ul>	<ul style="list-style-type: none"> <li>Unprecedented amount of heterogeneous data</li> <li>Large amount of user-generated data (text, video, image and audio)</li> <li>High speed generation rate</li> <li>Excessive usage of popular social networks</li> </ul>
13.	Jo et al.	2022	Korea	Telematics and Informatics		<ul style="list-style-type: none"> <li>Fact checking systems</li> <li>Provision of evidence-based facts</li> <li>Contingent evaluation methods</li> </ul>	<ul style="list-style-type: none"> <li>Infodemic</li> <li>Media reliability problem due to misinformation</li> </ul>

Table A1. Cont.

S.N.	Author	Year	Country	Journal	Relation of Big Data Analytics with Fake News Detection	Trending Approaches to Detect Fake News on Digital Media	Challenges for Constructing Quality Big Data to Detect Misinformation on Social Media
14.	Ebadi et al.	2020	United States	IEEE Transactions on Big Data		<ul style="list-style-type: none"> <li>Automated fact checking sites</li> <li>Classification based models</li> <li>Real-life fact-checking website</li> <li>Deep learning models</li> </ul>	<ul style="list-style-type: none"> <li>High speed of misinformation at social media sites</li> <li>Conspiracy and fake sites</li> </ul>
15.	Zrnec et al.	2022	Slovenia	Information Processing and Management		<ul style="list-style-type: none"> <li>IQ based tools</li> <li>Personality traits</li> </ul>	<ul style="list-style-type: none"> <li>High speed of fake news proliferation at social media</li> <li>Challenge of content evaluation</li> <li>Volume of data in the global data sphere</li> <li>Changing users' behaviors</li> <li>Overflowing of information resources</li> <li>Vast amount of content on digital media</li> </ul>
16.	Al-Rawi et al.	2018	Canada	Online Information Review			<ul style="list-style-type: none"> <li>Fake user accounts</li> <li>Easy spread of fake news on social media due to networked affordances</li> <li>Unmanageable spammy content</li> </ul>
17.	Qayyum et al.	2019	Pakistan	Cryptography and Security		<ul style="list-style-type: none"> <li>Generative machine learning</li> <li>Blockchain-based framework for fake news prevention</li> </ul>	Digitization of human life via social networking applications
18.	Jung et al.	2020	Germany	Big Data and Society	Big data analysis assists in uncovering digital fake news.	<ul style="list-style-type: none"> <li>Usage of official sources</li> <li>Deletion of rumor related content</li> </ul>	<ul style="list-style-type: none"> <li>Huge user-generated content</li> <li>Ideological polarization and a decreasing trust in traditional media</li> </ul>

Table A1. Cont.

S.N.	Author	Year	Country	Journal	Relation of Big Data Analytics with Fake News Detection	Trending Approaches to Detect Fake News on Digital Media	Challenges for Constructing Quality Big Data to Detect Misinformation on Social Media
19.	Kozik et al.	2022	Poland	Journal of Computational Science		<ul style="list-style-type: none"> <li>Text analysis to detect fake news and disinformation</li> <li>Artificial neural network</li> <li>Data classification</li> </ul>	
20.	Meesad	2021	Singapore	SN Computer Science		<ul style="list-style-type: none"> <li>Machine learning</li> <li>Effective information retrieval</li> </ul>	
21.	Liu	2019	USA	Journal of Services Marketing		Artificial intelligence tools	<ul style="list-style-type: none"> <li>Social bots significantly contribute fake information.</li> <li>Manipulation of facts via personal emotions</li> <li>Unchecked user generated content</li> </ul>
22.	Lewis and Westlund	2015	USA	Digital Journalism	Big data analytics and fake news detection are positively correlated with each other.		
23.	Veglis and Maniou	2018	Greece	KOME – An International Journal of Pure Communication Inquiry			<ul style="list-style-type: none"> <li>Digital journalism</li> <li>Sensational news for increasing rating</li> <li>Fast reach of online content</li> <li>Negative role of journalists, YouTubers</li> <li>Reshaping of the media landscape</li> </ul>
24.	Huckle and White	2017	United Kingdom	Big Data		Blockchain-based applications	<ul style="list-style-type: none"> <li>Diverse sources of conflicting information</li> </ul>
25.	Marquez et al.	2019	Spain	International Journal of Information Management		<ul style="list-style-type: none"> <li>Social media data analysis</li> <li>Textual review</li> <li>Big data architectures</li> <li>Machine learning</li> </ul>	The industry 4.0 is generating more data than ever before in the history of humanity.

Table A1. Cont.

S.N.	Author	Year	Country	Journal	Relation of Big Data Analytics with Fake News Detection	Trending Approaches to Detect Fake News on Digital Media	Challenges for Constructing Quality Big Data to Detect Misinformation on Social Media
26.	Bates et al.	2018	United States	Health Policy and Technology	Big data improves accuracy in health-related information.		
27.	Olmedilla et al.	2016	Spain	Computer Standards and Interfaces	Big data assists in detecting accurate information from online user-generated content.	Effective web-crawler	Huge amount of contextual data
28.	Shu et al.	2020	United States	Big Data		Computational solutions	<ul style="list-style-type: none"> <li>• Wide dissemination of fake news</li> <li>• Lack of comprehensive and community-driven fake news data sets</li> </ul>
29.	Awan et al.	2021	Pakistan	Int. J. Computer Applications in Technology			<ul style="list-style-type: none"> <li>• Fake profile trends</li> <li>• Security issues</li> </ul>
30.	Raza and Ding	2022	Canada	International Journal of Data Science and Analytics	Big data sets prove useful in fake news identification.	Social contexts to detect fake news	<ul style="list-style-type: none"> <li>• Difficult detection of fake news at early stage</li> <li>• Shortage of labelled data</li> </ul>
31.	Kauffmann et al.	2020	Spain	Industrial Marketing Management	Big data transformed into valuable information detects fake news.	<ul style="list-style-type: none"> <li>• Natural language processing technology</li> <li>• Sentiments analysis tools</li> <li>• Automatic fake news detectors</li> </ul>	
32.	King and Wang	2021	United States	International Journal of Information Management	Big data-driven approach finds out validity of online posted news.		
33.	Hassani et al.	2020	Iran	Big Data and Cognitive Computing	Text mining in big data analytics is a powerful tool against fake news on digital media.		

Table A1. Cont.

S.N.	Author	Year	Country	Journal	Relation of Big Data Analytics with Fake News Detection	Trending Approaches to Detect Fake News on Digital Media	Challenges for Constructing Quality Big Data to Detect Misinformation on Social Media
34.	Thota et al.	2018	United States	SMU Data Science Review		<ul style="list-style-type: none"> <li>• Deep learning architectures</li> <li>• Neural network</li> </ul>	
35.	Ahmad et al.	2020	Pakistan	Complexity		Machine learning ensemble approach	Rapid adoption of social media platforms
36.	Monti et al.	2019	United Kingdom	Social and Information Networks			Forming propagation patterns could be harnessed for the automatic fake news detection.
37.	Sahoo and Gupta	2021	India	Applied Soft Computing Journal		<ul style="list-style-type: none"> <li>• Identification of users' profiles</li> <li>• Automatic fake news detection approach in chrome environment</li> </ul>	
38.	Sharma et al.	2020	India	International Journal of Engineering Research & Technology		<ul style="list-style-type: none"> <li>• Artificial intelligence</li> <li>• Natural language processing</li> <li>• Machine learning</li> </ul>	Biased opinions
39.	Aslam et al.	2021	Saudi Arabia	Complexity	Ensemble-based deep learning model to classify news as fake or real using LIAR dataset		Diffusion of low-quality news in social media
40.	Chauhan and Palivela	2021	India	International Journal of Information Management Data Insights		<ul style="list-style-type: none"> <li>• Neural network</li> <li>• Deep learning-based approach</li> </ul>	



**Table A1.** *Cont.*

S.N.	Author	Year	Country	Journal	Relation of Big Data Analytics with Fake News Detection	Trending Approaches to Detect Fake News on Digital Media	Challenges for Constructing Quality Big Data to Detect Misinformation on Social Media
41.	Jiang et al.	2022	China	Information Processing and Management		Machine learning and deep learning methods	
42.	Galli et al.	2022	Italy	Journal of Intelligent Information Systems		<ul style="list-style-type: none"> <li>• Real-world datasets</li> <li>• Deep learning techniques</li> </ul>	Huge volumes of fake news posted by malicious users

## References

- Allcott, H.; Gentzkow, M. Social media and fake news in 2016 election. *J. Econ. Perspect.* **2017**, *31*, 211–236. [\[CrossRef\]](#)
- Golbeck, J.; Mauriello, M.; Auxier, B.; Bhanushali, K.H.; Bonk, C.; Bouzaghane, M.A.; Visnansky, G. Fake news vs satire: A dataset and analysis. In Proceedings of the 10th ACM Conference on Web Science, Amsterdam, The Netherlands, 27–30 May 2018; pp. 17–21.
- Al-Rawi, A.; Groshek, J.; Zhang, L. What the fake? Assessing the extent of networked political spamming and bots in the propagation of fakenews on Twitter. *Online Inf. Rev.* **2018**, *43*, 53–71. [\[CrossRef\]](#)
- Veglis, A.; Maniou, T.A. The mediated data model of communication flow: Big data and data journalism. *KOME Int. J. Pure Commun. Inq.* **2018**, *6*, 32–43. [\[CrossRef\]](#)
- Huckle, S.; White, M. Fake news: A technological approach to proving the origins of content, using blockchains. *Big Data* **2017**, *5*, 356–371. [\[CrossRef\]](#)
- Khan, J.Y.; Khondaker, M.d.T.I.; Afroz, S.; Uddin, G.; Iqbal, A. A benchmark study of machine learning models for online fake news detection. *Mach. Learn. Appl.* **2021**, *4*, 100032. [\[CrossRef\]](#)
- Qayyum, A.; Qadir, J.; Janjua, M.U.; Sher, F. Using blockchain to rein in the new post-truth world and check the spread of fake news. *IT Prof.* **2019**, *21*, 16–24.
- Jung, A.; Ross, B.; Stieglitz, S. Caution: Rumors ahead—A case study on the debunking of false information on twitter. *Big Data Soc.* **2020**, *7*, 1–15. [\[CrossRef\]](#)
- Mahabub, A. A robust technique of fake news detection using ensemble voting classifier and comparison with other classifiers. *SN Appl. Sci.* **2020**, *2*, 525. [\[CrossRef\]](#)
- Mayer-Schönberger, V.; Cukier, K. *Big Data: A Revolution That Will Transform How We Live, Work, and Think*; Houghton Mifflin Harcourt: Boston, MA, USA, 2013.
- Tan, W.; Blake, M.B.; Saleh, I.; Dustdar, S. Social-network-sourced big data analytics. *IEEE Internet Comput.* **2013**, *17*, 62–69.
- Olmedilla, M.; Martínez-Torres, M.R.; Toral, S.L. Harvesting big data in social science: A methodological approach for collecting online user-generated content. *Comput. Stand. Interfaces* **2016**, *46*, 79–87. [\[CrossRef\]](#)
- Marquez, J.L.J.; Gonzalez-Carrasco, I.; Lopez-Cuadrado, J.L.; Ruiz-Mezcua, B. Towards a big data framework for analyzing social media content. *Int. J. Inf. Manag.* **2019**, *44*, 1–12. [\[CrossRef\]](#)
- Hassani, H.; Beneki, C.; Unger, S.; Mazinani, M.T.; Yeganegi, M.R. Text mining in big data analytics. *Big Data Cogn. Comput.* **2020**, *4*, 1–34. [\[CrossRef\]](#)
- Bates, D.W.; Heitmueller, A.; Kakad, M.; Saria, S. Why policymakers should care about big data in healthcare. *Health Policy Technol.* **2018**, *7*, 211–216. [\[CrossRef\]](#)
- Torabi, A.F.; Taboada, M. Big data and quality data for fake news and misinformation detection. *Big Data Soc.* **2019**, *6*, 1–14. [\[CrossRef\]](#)
- Kauffmann, E.; Peral, J.; Gil, D.; Ferrández, A.; Sellers, R.; Mora, H. A framework for big data analytics in commercial social networks: A case study on sentiment analysis and fake review detection for marketing decision-making. *Ind. Mark. Manag.* **2020**, *90*, 523–537. [\[CrossRef\]](#)
- King, K.K.; Wang, B. Diffusion of real versus misinformation during a crisis event: A big data-driven approach. *Int. J. Inf. Manag.* **2021**, in press. [\[CrossRef\]](#)
- Ebadi, N.; Jozani, M.; Choo, K.R.; Rad, P. A memory network information retrieval model for identification of news misinformation. *IEEE Trans. Big Data* **2020**, *8*, 1358–1370. [\[CrossRef\]](#)
- Zrnec, A.; Pozenel, M.; Lavbic, D. Users' ability to perceive misinformation: An information quality assessment approach. *Inf. Process. Manag.* **2022**, *59*, 102739. [\[CrossRef\]](#)
- Supriyanto, E.E.; Bakti, I.S.; Furqon, M. The role of big data in the implementation of distance learning. *Paedagogia* **2021**, *12*, 61–68.
- Darwiesh, A.; Alghamdi, M.; El-Baz, A.H.; Elhoseny, M. Social media big data analysis: Towards enhancing competitiveness of firms in a post-pandemic world. *J. Healthc. Eng.* **2022**, *2022*, 6967158. [\[CrossRef\]](#)
- Thota, A.; Tilak, P.; Ahluwalia, S.; Lohia, N. Fake news detection: A deep learning approach. *SMU Data Sci. Rev.* **2018**, *1*, 10.
- Ahmad, I.; Yousaf, M.; Yousaf, S.; Ahmad, M.O. Fake news detection using machine learning ensemble methods. *Complexity* **2020**, *1*, 8885861. [\[CrossRef\]](#)
- Monti, F.; Frasca, F.; Eynard, D.; Mannion, D.; Bronstein, M.M. Fake news detection on social media using geometric deep learning. *Soc. Inf. Netw.* **2019**, *1*, 1–15.
- Sahoo, S.R.; Gupta, B.B. Multiple features based approach for automatic fake news detection on social networks using deep learning. *Appl. Soft Comput.* **2021**, *100*, 106983. [\[CrossRef\]](#)
- Sharma, U.; Saran, S.; Patil, S.M. Fake news detection using machine learning algorithms. *Int. J. Creat. Res. Thoughts (IJCRT)* **2020**, *8*, 509–518.
- Aslam, N.; Ullah Khan, I.; Alotaibi, F.S.; Aldaej, L.A.; Aldubaikil, A.K. Fake detect: A deep learning ensemble model for fake news detection. *Complexity* **2021**, *1*, 5557784. [\[CrossRef\]](#)
- Chauhan, T.; Palivela, H. Optimization and improvement of fake news detection using deep learning approaches for societal benefit. *Int. J. Inf. Manag. Data Insights* **2021**, *1*, 100051. [\[CrossRef\]](#)
- Vyas, P.; Liu, J.; El-Gayar, O.F. Fake News Detection on the Web: An LSTM-based Approach. In Proceedings of the AMCIS 2021, Digital Innovation and Entrepreneurship, Virtual, 9–13 August 2021; Volume 5.

31. Jiang, G.; Liu, S.; Zhao, Y.; Sun, Y.; Zhang, M. Fake news detection via knowledgeable prompt learning. *Inf. Processing Manag.* **2022**, *59*, 103029. [\[CrossRef\]](#)
32. Galli, A.; Masciari, E.; Moscato, V.; Sperli, G. A comprehensive benchmark for fake news detection. *J. Intell. Inf. Syst.* **2022**, *59*, 237–261. [\[CrossRef\]](#)
33. Nakamura, K.; Levy, S.; Wang, W.Y. A new multimodal benchmark dataset for fine-grained fake news detection. *Comput. Lang.* **2020**, *1*, 1–9.
34. Ianni, M.; Masciari, E.; Sperli, G. A survey of big data dimensions vs social networks analysis. *J. Intell. Inf. Syst.* **2020**, *57*, 73–100. [\[CrossRef\]](#)
35. Jo, H.; Park, S.; Shin, D.; Shin, J.; Lee, C. Estimating cost of fighting against fake news during catastrophic situations. *Telemat. Inform.* **2022**, *66*, 101734. [\[CrossRef\]](#)
36. Liu, X. A big data approach to examining social bots on twitter. *J. Serv. Mark.* **2019**, *33*, 369–379. [\[CrossRef\]](#)
37. Kozik, R.; Kula, S.; Choras, M.; Woźniak, M. Technical solution to counter potential crime: Text analysis to detect fake news and disinformation. *J. Comput. Sci.* **2022**, *60*, 101576. [\[CrossRef\]](#)
38. Meesad, P. Thai fake news detection based on information retrieval, natural language processing and machine learning. *SN Comput. Sci.* **2021**, *2*, 425. [\[CrossRef\]](#)
39. Raza, S.; Ding, C. Fake news detection based on news content and social contexts: A transformer-based approach. *Int. J. Data Sci. Anal.* **2022**, *13*, 335–362. [\[CrossRef\]](#)
40. Moher, D.; Shamseer, L.; Clarke, M.; Ghersi, D.; Liberati, A.; Petticrew, M.; Shekelle, P.; Stewart, L.A. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Syst. Rev.* **2015**, *4*, 1. [\[CrossRef\]](#)
41. Shahzad, K.; Khan, S.A. Factors affecting the adoption of integrated semantic digital libraries (SDLs): A systematic review. *Library Hi Tech* **2022**. *ahead of print*. [\[CrossRef\]](#)
42. Lewis, S.C.; Westlund, O. Big data and journalism: Epistemology, expertise, economics, and ethics. *Digit. J.* **2015**, *3*, 447–466. [\[CrossRef\]](#)
43. Guo, L.; Vargo, C.J. Global intermedia agenda setting: A big data analysis of international news flow. *J. Commun.* **2017**, *67*, 499–520. [\[CrossRef\]](#)
44. Murayama, T. Dataset of fake news detection and fact verification: A survey. *ACM Comput. Surv.* **2021**, *1*, 1–33.
45. Vargo, C.J.; Guo, L.; Amazeen, M.A. The agenda-setting power of fake news: A big data analysis of the online media landscape from 2014 to 2016. *New Media Soc.* **2018**, *20*, 2028–2049. [\[CrossRef\]](#)
46. Shu, K.; Mahudeswaran, D.; Wang, S.; Lee, D.; Liu, H. Fakenewsnet: A data repository with news content, social context, and spatiotemporal information for studying fake news on social media. *Big Data* **2020**, *8*, 171–188. [\[CrossRef\]](#)
47. Baur, N.; Graeff, P.; Braunsch, L.; Schweia, M. The quality of big data: Development, problems, and possibilities of use of process-generated data in the digital age. *Hist. Soc. Res.* **2020**, *45*, 209–243.
48. Awan, M.J.; Khan, M.A.; Ansari, Z.K.; Yasin, A.; Shehzad, H.M.F. Fake profile recognition using big data analytics in social media platforms. *Int. J. Comput. Appl. Technol.* **2021**, *68*, 215–222. [\[CrossRef\]](#)